The Astrophysical Journal, 632:137–156, 2005 October 10

NOAO FUNDAMENTAL PLANE SURVEY. II. AGE AND METALLICITY ALONG THE RED SEQUENCE FROM LINE-STRENGTH DATA

Jenica E. Nelan,[1] Russell J. Smith,[2] Michael J. Hudson,[2] Gary A. Wegner,[1] John R. Lucey,[3] Stephen A. W. Moore,[3] Stephen J. Quinney,[3] and Nicholas B. Suntzeff[4]

Received 2004 November 22; accepted 2005 May 15

ABSTRACT

We present spectroscopic line-strength data for 4097 red-sequence galaxies in 93 low-redshift galaxy clusters and use these to investigate variations in average stellar populations as a function of galaxy mass. Our analysis includes an improved treatment of nebular emission contamination, which affects ∼10% of the sample galaxies. Using the stellar population models of D. Thomas and collaborators, we simultaneously fit 12 observed line-strength–σ relations in terms of common underlying trends of age, [Z/H] (total metallicity), and [α/Fe] (α-element enhancement). We find that the observed line-strength–σ relations can be explained only if higher mass red-sequence galaxies are, on average, older, more metal-rich, and more α-enhanced than lower mass galaxies. Quantitatively, the scaling relations are age $\propto \sigma^{0.59\pm0.13}$, $Z/\mathrm{H} \propto \sigma^{0.53\pm0.08}$, and $\alpha/\mathrm{Fe} \propto \sigma^{0.31\pm0.06}$, where the errors reflect the range obtained using different subsets of indices. Our conclusions are not strongly dependent on which Balmer lines are used as age indicators. The derived age-σ relation is such that if the largest ($\sigma \sim 400$ km s^{-1}) galaxies formed their stars ∼13 Gyr ago, then the mean age of low-mass ($\sigma \sim 50$ km s^{-1}) objects is only ∼4 Gyr. The data also suggest a large spread in age at the low-mass end of the red sequence, with 68% of the galaxies having ages between 2 and 8 Gyr. We conclude that although the stars in giant red galaxies in clusters formed early, most of the galaxies at the faint end joined the red sequence only at recent epochs. This "downsizing" trend is in good qualitative agreement with observations of the red sequence at higher redshifts but is not predicted by semianalytic models of galaxy formation.

Subject headings: galaxies: clusters: general — galaxies: elliptical and lenticular, cD — galaxies: evolution — surveys

Online material: machine-readable tables

1. INTRODUCTION

Although early-type galaxies contain the bulk of the stellar mass in the low-redshift universe, their formation histories remain poorly understood. Early-type galaxies lie on a tight "red sequence" in the color-magnitude diagram (Sandage & Visvanathan 1978; Bower et al. 1992) and follow well-known dynamical scaling relations, such as the Faber-Jackson (Faber & Jackson 1976), the D_n-σ (Dressler 1987), and the fundamental plane (FP) relations (Djorgovski & Davis 1987). In addition to these scaling relations, elliptical galaxies also exhibit systematic correlations between spectroscopic absorption line strengths and velocity dispersion, σ. Especially well studied are the line strength centered on the magnesium triplet of absorption lines near 5175 Å with σ (Bender et al. 1993; Wegner et al. 1999; Kuntschner et al. 2001; Bernardi et al. 2003; Davies et al. 1987), as well as iron lines and the Balmer series (e.g., Mehlert et al. 2003).

The tightness of these scaling relations has generally been interpreted as evidence for coeval formation of early-type galaxies (Bower et al. 1992), with the slope of the color-magnitude relation arising from a mass-metallicity sequence. However, attempts to fit their spectra with stellar population models suggest that the formation of early-type galaxies may be more complicated and that there may be a spread in their ages (Worthey 1994; Thomas et al. 2003, hereafter TMB03; Trager et al. 2000b; Caldwell et al. 2003). Recent studies of early-type galaxies at high redshift indicate that the scatter and zero point of the FP relation vary with redshift (Wuyts et al. 2004) and may indicate that stellar population effects and age both play roles in determining the FP (van Dokkum & Ellis 2003).

A primary goal of the NOAO Fundamental Plane Survey (NFPS) is to address these questions by using early-type galaxies in nearby clusters ($z < 0.07$). The NFPS data set consists of photometry and spectroscopy for 5479 galaxies belonging to 93 clusters and is ideal for studying the evolution and properties of early-type galaxies in the cluster environment. The galaxies that have been selected for spectroscopic analysis are all cluster "red-sequence" galaxies. Therefore, they are mostly elliptical or S0 galaxies, but there is no explicit selection by morphology.

The NFPS data are to be presented in a series of papers. Smith et al. (2004, hereafter Paper I) described the goals and selection of the survey, and that work contains redshifts and velocity dispersions. The purpose of this paper, the second in the series, is to correlate the line strengths extracted from our high-quality spectroscopic data with our velocity dispersions and use the resulting relations to investigate the broad trends of stellar age and metallicity along the mass sequence. In § 2 we present the measurements of the absorption-line strengths and their errors, and in § 3 we detail our new method of measuring emission lines in the Hβ line strength. In § 4 we describe how we chose which galaxies to include in our final sample. In § 5 we show the derived line-strength–σ relations for many of the line strengths,

[1] Department of Physics and Astronomy, Dartmouth College, 6127 Wilder Laboratory, Hanover, NH 03755; current address: Department of Astronomy, Yale University, P.O. Box 208101, New Haven, CT 06520; jnelan@astro.yale.edu.

[2] Department of Physics, University of Waterloo, Waterloo, ON N2L 3G1, Canada.

[3] Department of Physics, University of Durham, South Road, Durham DH1 3LE, UK.

[4] Cerro Tololo Inter-American Observatory, Casilla 603, La Serena, Chile.

and in § 6 we use stellar population models to derive global age, metallicity, and α-element enhancement ([α/Fe]) trends as a function of velocity dispersion. We compare our trends to those from other studies and to results from intermediate-redshift observations in § 7.

2. ABSORPTION-LINE MEASUREMENTS

2.1. *Spectroscopic Data*

Paper I contains detailed descriptions of the selection of the data for spectroscopic follow-up along with the observations and measurements of redshifts and velocity dispersions. To summarize, NFPS clusters were X-ray selected from the XBACS (Ebeling et al. 1996) and BCS (Ebeling et al. 1998) catalogs for imaging. Individual galaxies were selected for follow-up spectroscopic observing based on their position in the cluster color-magnitude diagram relative to the red-sequence ridgeline. Specifically, galaxies with $R < 17$ and $\Delta(B - R) > 0.2$ were chosen for spectroscopic observation. As noted above, there are no explicit morphological selection criteria.

NFPS spectroscopic observations were carried out using the Hydra multifiber spectrographs at the Cerro Tololo Inter-American Observatory (CTIO) Blanco 4 m telescope and the 3.5 m WIYN telescope at Kitt Peak. Approximately 50–70 galaxies were observed in each cluster. Spectral resolution for both CTIO and WIYN is 3 Å, with CTIO data sampled at 1.15 Å pixel^{-1} and WIYN data sampled at 1.4 Å pixel^{-1}. The median signal-to-noise ratio (S/N) for both WIYN and CTIO data is 22. Variance-weighted extraction of the spectra, along with cosmic-ray rejection, wavelength calibration, and sky subtraction, was performed using tasks in the HYDRA package in IRAF.

2.2. *Line-Strength Measurements*

The line-strength indices used in this paper are from the original Lick/IDS system (Burstein et al. 1984) and its extensions to high-order Balmer lines (Worthey & Ottaviani 1997), including Hβ^+ (González 1993). Note that the Hδ_A and Hδ_F line strengths from Worthey & Ottaviani (1997) differ slightly from the Hδ line strength used in, e.g., Balogh et al. (1999); the Hδ_A central bandpass is almost identical to their Hδ, but the red and blue continua differ by $\sim$10 Å on each side. Hδ_F is a much narrower line strength than that of Balogh et al. (1999).

Additionally, we defined and measured two new line strengths around the Hα line, which falls within the range of the WIYN spectra, for $\sim$700 galaxies with $cz \lesssim 13{,}000$ km s^{-1}. The two line strengths, summarized in Table 1, differ in their sensitivity to contamination from neighboring [N II] emission lines. Hα_F has a narrow definition, with a narrow red continuum bandpass between the [N II] λ6583 emission line and Hα. This is primarily designed to detect Hα emission where present. Hα_A has a wider red continuum, which is contaminated by N II emission when present. In the absence of emission, however, Hα_A should be the more reliable indicator of stellar Hα absorption.

The molecular TiO_1 and TiO_2 line strengths are only measured in $\sim$1000 of our galaxies, most of which are from our northern clusters since the wavelength range of the WIYN spectrograph extends further into the red than that at CTIO. Similarly, relatively few galaxies have Hα measurements for the same reason.

Prior to measuring line strengths, the galaxy spectra were corrected to an approximate relative flux scale by comparison with a model elliptical template from Kinney et al. (1996). However, because we did not observe Lick calibration stars, we have not attempted to correct our galaxy spectra in order to directly match to the Lick flux system. Because each line strength is defined using a pair of pseudocontinua bracketing the line feature, the line-strength indices are robust against differences in the Lick and NFPS response functions unless there is substantial relative curvature over the extent of a line strength. The most susceptible definitions are those with the widest extent, viz., the "molecular" line strengths $Mg_{1,2}$, $TiO_{1,2}$, and $CN_{1,2}$. Caution should be used when analyzing these line strengths, especially if the sample covers a substantial range in redshift. Note also that the NFPS spectra are unsuitable for the measurement of discontinuity line strengths such as D4000, since these are very sensitive to flux calibration. Moreover, our instrumental setup, especially at WIYN, has very poor response at $\lesssim$4000 Å.

TABLE 1
NFPS Hα Line Indices

Line Index	Bandpass (Å)	Pseudocontinua (Å)
Hα_F	6554.000–6568.000	6515.000–6540.000
	6554.000–6568.000	6568.000–6575.000
Hα_A	6554.000–6575.000	6515.000–6540.000
	6554.000–6575.000	6575.000–6585.000

To measure line strengths, we chose to use the program INDEXF (Cardiel, Gorgas, and Cenarro; see Cenarro et al. 2001) because of its careful calculation of line-strength errors, following Cardiel et al. (1998). The estimated error spectrum (incorporating Poisson errors and a contribution from sky subtraction noise) yields a "photometric" error, which is added in quadrature with the (Monte Carlo derived) errors due to the uncertainty in radial velocity, yielding a total error due to photon noise, $\epsilon_{\rm phot}$.

The line strengths are measured both at the $\sim$3 Å native resolution of the NFPS spectra and also at the $\sim$9 Å resolution of the Lick system. The galaxy spectra were broadened to the Lick resolution for each line strength following the resolution curve from Worthey & Ottaviani (1997). For the NFPS Hα line strengths that lie redward of the Lick spectral range, we smoothed by 10.5 Å FWHM. The Lick-resolution line strengths are suitable for comparison to synthesis models such as those of Worthey (1994), based on low-resolution stellar libraries. The full-resolution line strengths in principle retain more detailed information (at least for galaxies with small velocity broadening) and can be analyzed in comparison to higher resolution synthesis models, such as those of Vazdekis (1999).

2.3. *Velocity Broadening and Aperture Corrections*

Galaxy spectra are broadened by the line-of-sight velocities of their stars, generally resulting in a dilution in measured line strength with increasing velocity dispersion, σ. We followed the standard procedure of determining correction curves from artificially broadened K giant stellar spectra. The correction curve for the WIYN data is the average of the curves from nine stars, while the CTIO correction curve is the average of the curves from 12 stars. These corrections are very stable and therefore we applied a common correction curve to data from all of the observing runs. For our Lick-resolution galaxy spectra, the stars were first broadened to the Lick resolution before deriving the correction curves. Our corrections are very similar to those obtained by other groups (e.g., Poggianti et al. 2001). The velocity broadening corrections derived from different template stars agree within $\lesssim$10% for all indices except Hβ. However, in absolute terms the velocity broadening correction for Hβ is quite small ($\lesssim$20% of the index value).

The NFPS spectra sample the galaxy light within a fixed 2″ diameter. This angular scale samples different physical scales for galaxies at different angular diameter distances. Because galaxies have internal gradients, it is necessary to correct the raw data for such aperture effects. In Paper I we corrected the velocity dispersions according to the prescription of Jorgensen et al. (1995).

Aperture corrections can also be applied to line strengths to correct for increasing apparent galaxy size at higher redshift. We followed the formula

$$I_{\rm cor} = I_{\rm ap} + \Delta_{\rm ap}, \quad (1)$$

where $I_{\rm ap}$ is the uncorrected line-strength value measured through a 2″ aperture and $\Delta_{\rm ap}$ is given by

$$\Delta_{\rm ap} = -\kappa \log\left(\frac{d}{d_0}\right). \quad (2)$$

The line-strength gradient $\kappa = d\log(\text{index})/d\log(r_{\rm ap})$, and d/d_0 is a ratio of angular diameter distances. Here d_0 is a normalization factor defined as the angular diameter distance at $z = 0.05$. For each galaxy, d is the angular diameter distance corresponding to the cluster CMB-frame redshift; for noncluster galaxies, d is a function of the CMB-frame redshift of the galaxy itself. In calculating d, we adopted a cosmology with $\Omega_m = 0.3$ and $\Omega_\lambda = 0.7$. Thus, the physical diameter of the corrected aperture diameter is 1.37 h^{-1} kpc. Note that this correction is to a fixed metric diameter and not to a multiple of the effective radius, R_e. The gradients κ were compiled from other groups that either measured internal line-strength gradients at varying radii or extracted multiple apertures for a galaxy, and they are summarized in Table 2. In particular, gradients from either Kuntschner et al. (2002) or Proctor (2002) were used for most of our line strengths. Table 2 also lists the mean difference between the corrected and uncorrected line strengths for each line index, as well as the dispersion in this difference. The difference between the corrected and uncorrected line strengths varies by galaxy but is generally at the level of a few percent.

The aperture corrections in velocity dispersion and line strengths were not applied for purposes of comparisons in §§ 2.4 and 2.5. However, in § 5 and subsequent sections, the velocity dispersions and line strengths are corrected for aperture effects.

TABLE 2
APERTURE CORRECTIONS

Line Index	Units	κ	$\langle\Delta_{\rm ap}\rangle$	$\langle\Delta_{\rm ap}^2\rangle^{1/2}$
$H\delta_A$	Å	1.283	−0.071	0.172
$H\delta_F$	Å	0.351	−0.019	0.047
CN_1	mag	−0.058	0.003	0.008
Ca4227	Å	−0.121	0.007	0.016
$H\gamma_A$	Å	1.349	−0.075	0.181
$H\gamma_F$	Å	0.629	−0.035	0.084
Fe4383	Å	−1.484	0.082	0.199
Ca4455	Å	−0.226	0.013	0.030
Fe4531	Å	−0.544	0.030	0.073
Fe4668	Å	−2.580	0.143	0.346
$H\beta$	Å	0.000	0.000	0.000
Fe5015	Å	−1.346	0.075	0.180
Mg_1	mag	−0.040	0.003	0.007
Mg_2	mag	−0.066	0.005	0.011
Mg b	Å	−1.375	0.080	0.182
Fe5270	Å	−0.706	0.091	0.095
Fe5335	Å	−0.706	0.049	0.118
Fe5406	Å	−0.460	0.014	0.062
Fe5709	Å	−0.019	0.001	0.003
Fe5782	Å	−0.019	0.001	0.002
Na5895	Å	−0.045	0.006	0.006
TiO_1	mag	−0.016	0.002	0.003
TiO_2	mag	−0.011	0.001	0.001
$H\alpha_A$	Å	0.035	−0.008	0.003
$H\alpha_F$	Å	0.035	−0.008	0.004

2.4. *Interrun Comparisons*

As described above, we chose to measure our line strengths with INDEXF largely because of its method of calculating errors, which allows for noise in the spectra and errors in the radial velocities. However, given the possibility of systematic errors, it is useful to check the consistency of the error estimates by comparing repeat observations of the same galaxies.

Within the NFPS survey several "standard" equatorial clusters were reobserved on several runs and from each telescope. The primary goal of this was to ensure that velocity dispersions were on a consistent base system and to correct discrepant runs if necessary. In Paper I we showed that velocity dispersions could be calibrated to a systematic accuracy of ∼0.004 dex (thus limiting systematic errors in FP-derived distances to ≤1.5%).

As an example of the consistency of our line-strength measurements, we compare the Mg b instrumental resolution absorption line measurements in galaxies that were observed in multiple runs at the same telescope and also at both sites. In Figure 1 we show comparisons of galaxies observed in more than one CTIO or WIYN run.

While Poisson errors are the dominant source of error in the line-strength index measurements, results for the same galaxy may also differ due to other effects such as fiber position errors and variations in fiber transmission. Such errors will affect individual measurements in a way that is likely to be independent of the S/N. To allow for these effects, we include an "external error," $\epsilon_{\rm fib}$, assumed to be constant for all measurements (but of course different from one line strength to another), added in quadrature with the estimated measurement errors, so that the total random error $\epsilon_{\rm ran} = (\epsilon_{\rm phot}^2 + \epsilon_{\rm fib}^2)^{1/2}$. We determine the value of this external error by requiring a good χ^2 for repeated measurements of the galaxies in common. We find that for most indices, the external errors are typically much smaller than the typical random errors. Nevertheless, it is necessary to account for them since accurate characterization of measurement errors is important in § 6 when we attempt to determine the intrinsic population variation in line strength. For six indices ($Mg_{1,2}$, $TiO_{1,2}$, $H\alpha_A$, $H\alpha_F$), the external errors are slightly larger than the typical random error. We do not use any of these six indices to determine stellar population parameter gradients.

There is also the possibility of small systematic offsets in the line-strength measurements between runs, due, for example, to variations in seeing or to different instrumental configurations. To calculate additive corrections, $\Delta_{\rm run}$, needed to bring the line-strength measurements from different runs onto a common system, we use the repeat measurements and follow the same procedure as was used for velocity dispersions in Paper I. We find that in most cases (with the exception of the $H\alpha$ indices) the run corrections are smaller than the typical random errors and are very much smaller than the range spanned by the line-strength data. Nevertheless, we prefer to include these small corrections so as not to bias future bulk flow measurements derived from the FP relation that also incorporate these line strengths as additional parameters. These corrections are applied to the raw line strengths before the data are merged to yield a final set

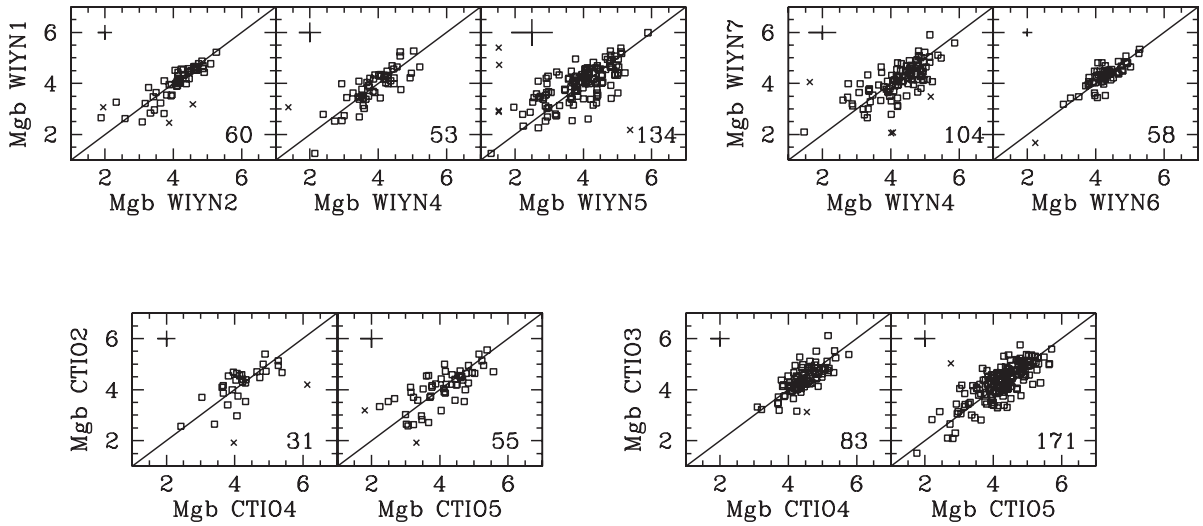

FIG. 1.—Overlapping observations between CTIO runs and WIYN runs. The line in each plot has a slope of unity, representing where galaxies should fall if there is no offset between multiple observations. Crosses indicate galaxies more than 4 σ away from the mean-weighted least-squares fit between the two axes. Average errors are shown in the upper left of the plot, and the number of galaxies is displayed in the lower right. Mg b is in units of Å and measured at full (instrumental) resolution.

TABLE 3
LINE-STRENGTH INDEX ERRORS

Line Index	$\langle I^2 \rangle^{1/2}$ [a]	$\epsilon_{\rm phot}$ [b]	$\epsilon_{\rm fib}$ [c]	$\epsilon_{\rm ran}$ [d]	$\langle \Delta^2_{\rm run} \rangle^{1/2}$ [e]
$H\delta_A$	2.14	0.91	0.58	0.99	0.30
$H\delta_F$	1.24	0.61	0.33	0.64	0.13
CN_1	0.06	0.02	0.02	0.03	0.01
CN_2	0.07	0.03	0.02	0.03	0.02
Ca4227	0.60	0.40	0.20	0.42	0.06
G4300	1.08	0.61	0.36	0.66	0.28
$H\gamma_A$	1.51	0.64	0.45	0.73	0.32
$H\gamma_F$	0.88	0.39	0.20	0.41	0.08
Fe4383	1.15	0.75	0.37	0.78	0.26
Ca4455	0.53	0.43	0.00	0.40	0.08
Fe4531	0.77	0.53	0.25	0.55	0.09
Fe4668	1.51	0.68	0.50	0.80	0.33
$H\beta$	0.60	0.27	0.11	0.28	0.05
$H\beta_p$	0.52	0.18	0.09	0.20	0.02
Fe5015	1.04	0.59	0.33	0.64	0.14
Mg_1	0.03	0.01	0.01	0.01	0.01
Mg_2	0.04	0.01	0.01	0.01	0.01
Mg b	0.66	0.26	0.12	0.27	0.03
Fe5270	0.44	0.28	0.11	0.28	0.04
Fe5335	0.50	0.34	0.11	0.34	0.05
Fe5406	0.35	0.25	0.11	0.26	0.04
Fe5709	0.27	0.20	0.10	0.21	0.02
Fe5782	0.27	0.17	0.10	0.19	0.05
Na5895	0.95	0.18	0.16	0.23	0.08
TiO_1	0.01	0.00	0.01	0.01	0.00
TiO_2	0.01	0.00	0.01	0.01	0.00
$H\alpha_A$	1.23	0.15	0.15	0.21	0.31
$H\alpha_F$	2.67	0.14	0.19	0.24	0.09

[a] The total dispersion over the NFPS in the measured index, arising from intrinsic population differences and measurement error.
[b] Median random error.
[c] Mean systematic error.
[d] Median total random error in merged sample.
[e] The rms correction over all runs.

of line-strength data for each galaxy. In Table 3, for each line strength we list the total dispersion in the measured line strengths for a given index, the typical photon error, $\epsilon_{\rm phot}$, the systematic error, $\epsilon_{\rm fib}$, the total random error, $\epsilon_{\rm ran}$, and the typical (rms) run correction, $\langle \Delta^2_{\rm run} \rangle^{1/2}$.

2.5. *Comparisons with Other Surveys*

We compare our line strengths with data from two other surveys: ~140 galaxies overlapping with the second data release from the Sloan Digital Sky Survey (Abazajian et al. 2004) and 33 Coma Cluster galaxies from Moore et al. (2002). The comparisons with Moore et al. (2002) are at the Lick resolution. SDSS line strengths available in their archive are at their instrumental resolution (2.4 Å), while our instrumental resolution line strengths are at 3 Å. Using the ratio of our Lick-resolution line strengths to our full-resolution line strengths, we "scaled" the SDSS line strengths to 3 Å resolution. The correction to each line strength is typically very small (~2%). None of the data sets have aperture corrections applied for these comparisons.

In Table 4 we compare our measurements with those from other surveys, for several line strengths in common. The mean offsets are defined as $\Delta = \langle I_{\rm NFPS} - I_{\rm other} \rangle$, where $I_{\rm NFPS}$ is the NFPS line strength and $I_{\rm other}$ is the SDSS or Moore et al. (2002) line strength, and the rms is the standard deviation of the differences.

We plot the NFPS line strengths against those of Moore et al. (2002) in Figure 2. Our data are in very good agreement with those of Moore et al. (2002); with the exception of Mg_1, which, as noted above, is susceptible to flux calibration error, there is no evidence for systematic offsets. The χ^2 values of the comparisons are acceptable for all line strengths except for Fe4668, Fe5270, and Mg_2. This suggests that, in most cases, both NFPS and Moore et al. (2002) errors are reasonable.

The comparison with SDSS is shown in Figure 3. Note that for line strengths in the blue, the quoted SDSS errors are smaller, whereas for line strengths in the red, the NFPS errors are typically smaller. There is evidence for small but significant offsets for most of the line strengths, as indicated in Table 4. Furthermore, for most line strengths, the reduced χ^2 values are larger than unity, suggesting that the errors in either SDSS or NFPS (or both) are underestimated. Overall, however, based on the consistency between our measurements and those of Moore et al.

TABLE 4
LINE INDEX OFFSETS

Survey	Line Index	Units	N_{gal}	Mean Offset	rms
Moore et al. (2002).....	Fe4668	Å	34	0.236 ± 0.174	0.694
	Hβ	Å	34	−0.035 ± 0.034	0.625
	Fe5015	Å	34	0.186 ± 0.119	0.769
	Mg_1	mag	34	0.017 ± 0.002	0.019
	Mg_2	mag	34	−0.005 ± 0.003	0.030
	Mg *b*	Å	34	−0.079 ± 0.036	0.473
	Fe5270	Å	34	−0.039 ± 0.058	0.420
	Fe5335	Å	34	0.063 ± 0.049	0.340
SDSS.....	Ca4227	Å	141	0.102 ± 0.039	0.468
	Fe4383	Å	141	0.261 ± 0.096	1.139
	Ca4455	Å	141	0.271 ± 0.046	0.543
	Fe4531	Å	141	0.231 ± 0.092	1.090
	Fe4668	Å	141	−0.046 ± 0.115	1.362
	Hβ	Å	141	0.021 ± 0.045	0.540
	Fe5015	Å	140	0.601 ± 0.078	0.920
	Mg *b*	Å	136	0.126 ± 0.039	0.455
	Fe5270	Å	64	0.220 ± 0.066	0.529
	Fe5335	Å	102	0.406 ± 0.054	0.547
	Fe5406	Å	135	0.208 ± 0.037	0.431
	Fe5709	Å	129	0.095 ± 0.032	0.366
	Fe5782	Å	84	0.021 ± 0.031	0.287
	Na5895	Å	71	0.083 ± 0.052	0.437

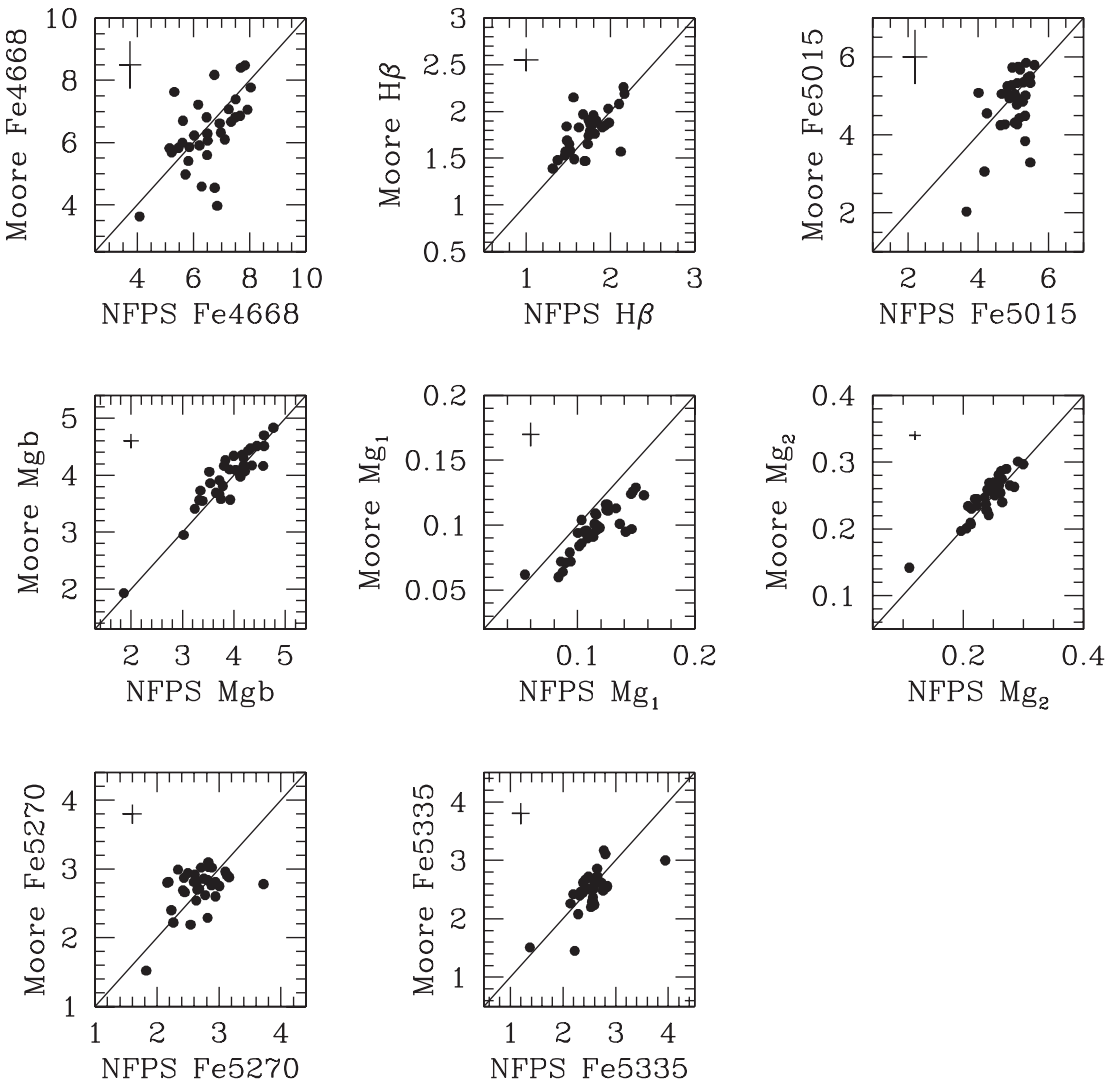

FIG. 2.—NFPS line strengths for several line indices plotted against those from Moore et al. (2002). Solid lines with a slope of unity indicate exact agreement between the two sets of measurements. All measurements are at Lick resolution. The average NFPS and Moore et al. (2002) line-strength error is shown in the upper left corner of each plot.

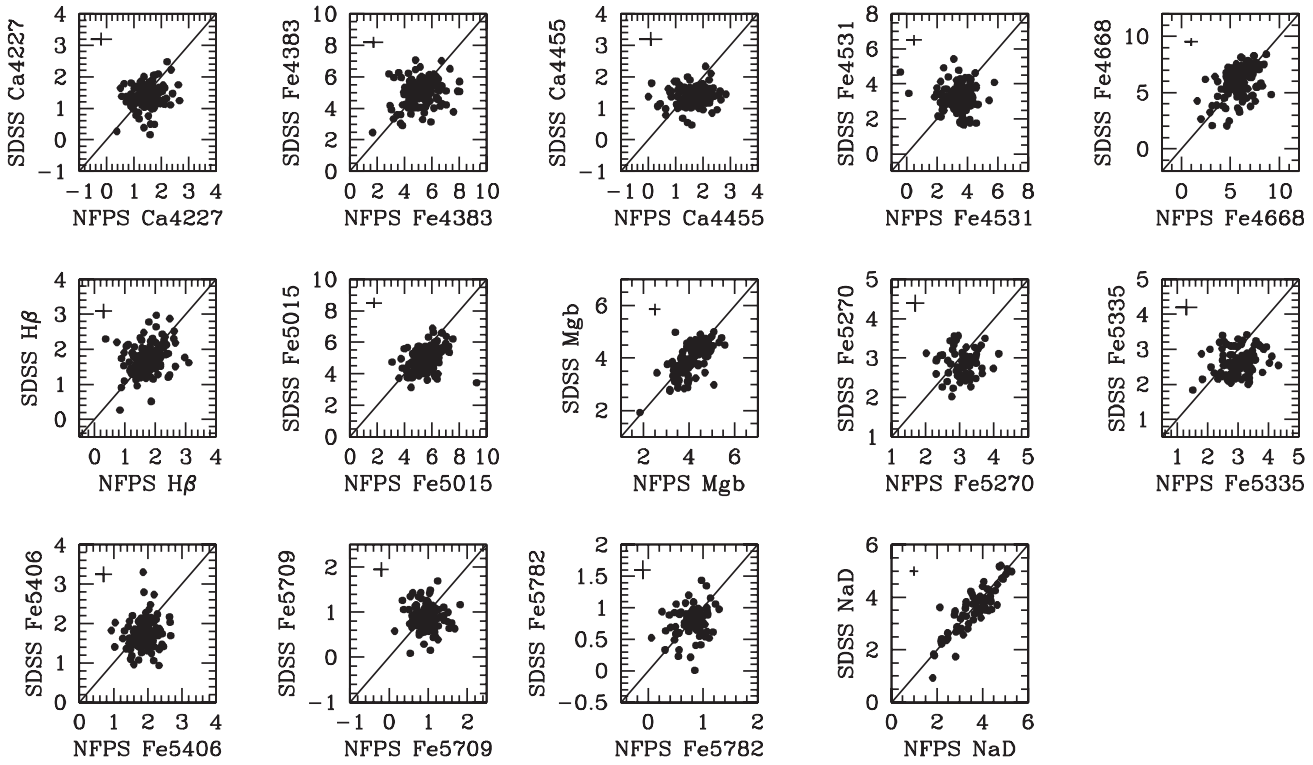

FIG. 3.—Plots of NFPS line strengths at instrumental resolution against SDSS line strengths "scaled" to the same NFPS resolution. Solid lines with a slope of unity indicate exact agreement between the two sets of measurements. The average NFPS and SDSS line-strength error is shown in the upper left corner of each plot.

(2002), we conclude that our error estimates, which include systematic effects estimated from repeat observations, are realistic.

3. NEBULAR EMISSION MEASUREMENTS

While elliptical galaxy spectra are broadly characterized as being dominated by absorption lines, it has long been realized that many systems also show nebular emission lines (Herbig & Mayall 1957; Phillips et al. 1986; González 1993; Goudfrooij et al. 1994).

A particular difficulty in studying integrated stellar populations is that nebular Hβ emission acts to "fill in" the stellar absorption line, driving derived age measurements toward older values. Disentangling the absorption and emission components of Hβ is therefore vital if this line strength is to be used to constrain star formation histories. Since emission at Hβ is often coincident with emission in [O III] $\lambda\lambda$4959, 5007], a standard approach is to establish a correction based on the more easily measured [O III] lines, usually assuming a constant correction factor of H$\beta = 0.6\times$[O III λ5007] (Trager et al. 2000a). In this section we discuss an alternative method that aims to distinguish whether galaxies are likely affected by emission by fitting the spectra with absorption template models.

3.1. *Method*

Since galaxy spectra contain substantial structure from unresolved stellar absorption lines, weak [O III] lines can be very difficult to distinguish without first removing the stellar component. This can be done by subtracting/dividing a "matched" spectrum for similar galaxies of similar stellar population properties but free from emission (Goudfrooij et al. 1994) or by subtracting/dividing a model stellar continuum (Kuntschner et al. 2002). In the latter case it is necessary to exclude the region around Mg *b* where current synthetic spectra (such as those of Vazdekis 1999) provide a poor match to observed elliptical galaxies due to non-solar abundance ratios. This limits the nebular lines that can be measured, excluding such potentially interesting features as N I λ5199, but the method has the advantage of being easily automated when the redshift and velocity dispersion are known.

Our measurements of Hβ and [O III] emission are made on the spectra after dividing by the best-fitting Vazdekis (1999) model, computed over the 4800–5100 Å spectral range. The models span a range of metallicity $-0.7 < [\mathrm{Fe/H}] < +0.2$ and age 1.0 Gyr $< t <$ 17.4 Gyr and assume solar abundance ratios and a power-law initial mass function (IMF) with $x = 1.3$. The best-fit model is computed by comparing the model and observed spectra after first shifting and broadening the model to match the redshift and velocity dispersion of the galaxy, as determined in Paper I. After division by the best-fitting model, a low-order continuum is divided out to remove long-wavelength baseline variations. Equivalent widths are measured directly on the divided spectra (i.e., without assuming an emission-line profile), and errors estimated based on the noise in the "line-free" regions. Figure 4 shows some illustrative examples of this continuum removal method.

It is clear that the model-fitting procedure itself can be biased by the presence of Hβ emission. Specifically, for galaxies with emission, a model with weaker stellar Hβ can be fitted, yielding an underestimate of the nebular Hβ. We have investigated this effect through simulations, adding emission lines to model population spectra of varying ages and velocity broadening. The recovered Hβ are biased low by 10%–25%, depending on the underlying spectrum. In spectra of high velocity dispersion, the broad wings of the absorption spectra help to distinguish the narrow nebular emission lines. In older populations, the bias is somewhat reduced because there is little flexibility to push the fit to even older models. Thus, the worst cases (where emission Hβ is underestimated by 25%) are for low-σ young stellar population spectra. Despite this bias, the wings of the Hβ feature

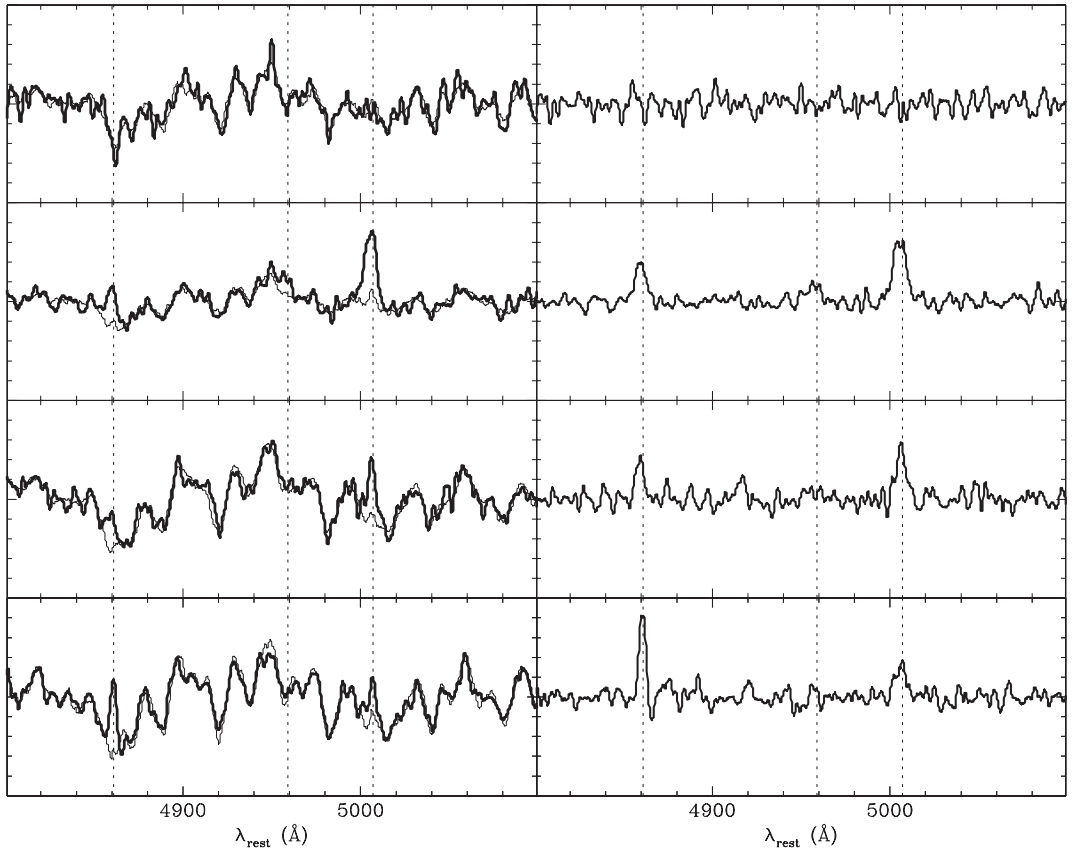

FIG. 4.—Some illustrative examples of the emission-line measurements. In the left-hand panels are shown the observed spectrum (*thick line*) and the best-fit Vazdekis (1999) model (*thin line*). The ratio spectrum, from which the line measurements are derived, is shown in the corresponding right-hand panel. Vertical dotted lines indicate the expected location of the Hβ and [O III] lines. At the top, a spectrum with no measurable emission is shown; in the lower three panels, spectra with weak to moderate emission. Note the substantial range of Hβ–to–[O III] ratios among these examples.

are essential to separate the broad stellar and narrow nebular contributions.

3.2. *Results*

The above method was applied to 4964 galaxies with redshift and velocity dispersion data. (For this purpose we allowed the use of velocity dispersions from lower S/N spectra not reported in Paper I.) Of these, significant [O III λ5007] emission is detected in 589 galaxies, EW([O III λ4959]) in 154 galaxies, and EW(Hβ) in 633 galaxies (all 3 σ detection limits).

The correlations between EW([O III λ5007]), EW(Hβ), and σ are presented in Figure 5. From Figure 5*a* it is clear that while the presence of O III is indeed an indicator for Hβ emission, no single ratio between the lines is appropriate for all galaxies. In particular, there is a substantial population of objects with moderate to strong Hβ emission but with little or no [O III]. Plotting the emission lines versus velocity dispersion (Figs. 5*b* and 5*c*) shows that these objects are overwhelmingly of low mass ($\log \sigma \lesssim 2.0$). Thus, for low-σ objects, the classical factor 0.6 undercorrects for emission contamination, while for high-σ

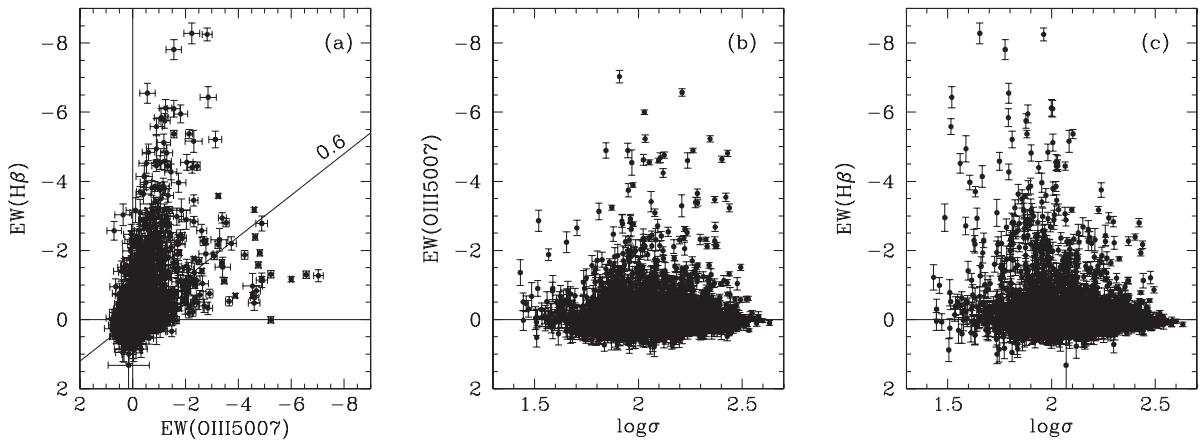

FIG. 5.—Emission measurements at [O III] λ5007 and Hβ as a function of velocity dispersion, σ.

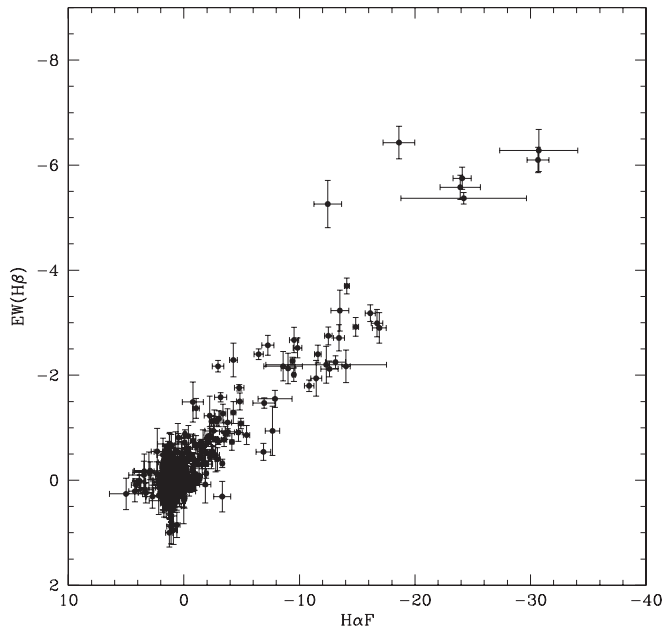

FIG. 6.—Emission-line measurements, vs. the $H\alpha_F$ line strength (see § 3.2). Units are in Å.

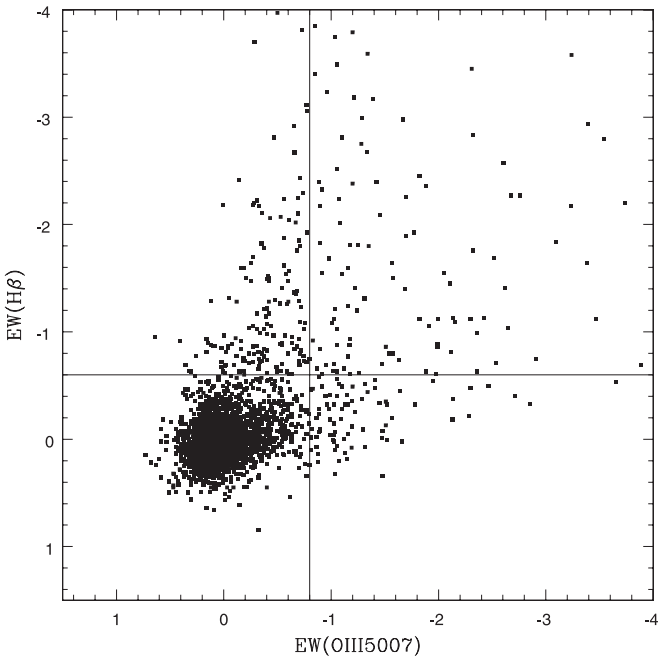

FIG. 7.—Equivalent widths of O III λ5007 and Hβ emission lines. Solid lines are lines of constant emission. Points above the horizontal line or right of the vertical line are galaxies excluded from our CULL data set. Units are in Å.

objects, the factor overcorrects. Note that the objects with large Hβ emission are those for which we expect in fact to *underestimate* the measurement, due to the bias described above.

A better predictor of emission at Hβ is the emission at Hα (Caldwell et al. 2003). As described in § 2.2, $H\alpha_F$ line-strength measurements are only available for a subset of ~700 galaxies, but these can provide a consistency check on Hβ emission, as shown in Figure 6. The $H\alpha_F$ line strength is constructed so as to be insensitive to the neighboring [N II] lines. While $H\alpha_F$ does of course have some contribution from stellar absorption, the ratio between the nebular emission lines is such that emission, when present, usually dominates at Hα, even if it is weak at Hβ. The excellent correlation in Figure 6 suggests that the emission Hβ measurements are indeed at least a reasonable proxy for Hα. The Hα/Hβ ratio ranges from 2.9 to 5.9 (68% range), with a median of 4.5, similar to the typical range of 2–6 for early-type spiral galaxies (Stasińska et al. 2004).

3.3. *Emission-Line Rejection Criteria*

For galaxies with excess emission in the Hβ and O III $\lambda\lambda$4959 and 5007 lines, the Hβ absorption line index is contaminated and hence the derived stellar properties are biased. As described above, a constant correction factor is incorrect, even in a statistical sense, since the true ratios vary widely and *systematically* with velocity dispersion. Therefore, instead of using the constant correction factor based only on [O III], we prefer simply to exclude galaxies with emission at either Hβ or [O III λ5007]. In Figure 7 we plot the emission measurements of Hβ against [O III λ5007] and indicate the cuts made to exclude high-emission galaxies. Only galaxies with EW(Hβ) > -0.6 and EW(O III) > -0.8 are included in our sample of all line strengths used for analysis in the next two sections.

4. DATA PRESENTATION AND SUMMARY

Tables 5 and 6 present absorption- and emission-line strengths for all galaxies observed, including those with emission. Hereafter we refer to this as the ALL data set. Table 5 gives line strengths and errors at Lick resolution for galaxies with S/N > 15 Å^{-1}, at 5000–5500 Å. The line strengths in the table are corrected for velocity broadening but have not been corrected for aperture effects. Additionally, the O III and Hβ emission equivalent widths with their common error are listed (indicated by W), along with the heliocentric redshift ($cz_\odot$), velocity dispersion (σ), cluster identification, and ratio of the angular diameter distances based on CMB-frame redshifts. Table 6 presents the full-resolution data in analogous form. The auxiliary parameters ($cz_\odot$, etc.) are duplicated here for convenience.

The final galaxy sample is culled for the analysis of the line-strength–σ trends and age and metallicity scaling relations in §§ 5 and 6.

5. LINE-STRENGTH–σ RELATIONS

The overall trends of line strengths with velocity dispersion reflect changes in the characteristic stellar populations as a function of galaxy mass. In this section we describe the line-strength–σ relations observed in the NFPS data and compare, where possible, to previous results.

5.1. *Sample*

In the remainder of this paper we analyze a restricted sample of the NFPS galaxies, hereafter referred to as the CULL data set. This is defined by applying the following cuts:

1. Galaxies must belong to one of the clusters in the NFPS cluster sample, as determined in Paper I. This rejects objects in the field and in background groups or clusters.
2. Galaxies must be free of emission lines as determined in this paper, having emission weaker than 0.8 Å in O III λ5007 *and* weaker than 0.6 Å in Hβ.

Moreover, in the CULL data set, the velocity dispersions and line strengths are corrected for aperture effects as described in § 2.3.

TABLE 5
Lick Resolution Line-Strength Indices and Associated Data

Galaxy ID			$H\delta_F$	$H\delta_A$	CN_1	CN_2	Ca4227	G4300	$H\gamma_F$
W(O III)	$W(H\beta)$	$\varepsilon(W)$	$H\gamma_A$	Fe4383	Ca4455	Fe4531	Fe4668	$H\beta$	$H\beta^+$
$cz_\odot$	$\log\sigma$		Fe5015	Mg b	Mg_1	Mg_2	Fe5270	Fe5335	Fe5406
Cluster	$\log(d/d_0)$		Fe5709	Fe5782	Na5895	TiO_1	TiO_2	$H\alpha_A$	$H\alpha_F$
NFP J001006.3−284623			1.687 ± 0.748	−1.355 ± 1.495	0.0266 ± 0.0416	0.0574 ± 0.0382	0.016 ± 0.698	6.174 ± 0.881	−0.648 ± 0.673
−0.06	−0.35	0.22	−4.244 ± 1.202	4.579 ± 1.062	0.674 ± 0.582	3.540 ± 0.852	4.268 ± 1.415	2.068 ± 0.459	2.342 ± 0.348
18989	1.877 ± 0.050		4.472 ± 0.921	3.051 ± 0.442	0.0435 ± 0.0133	0.1643 ± 0.0170	...	2.884 ± 0.504	1.004 ± 0.411
A2734	0.080		1.499 ± 0.365	...	...	...	...	...	...
NFP J001010.7−285020			0.580 ± 0.553	−1.020 ± 1.001	0.0192 ± 0.0306	0.0446 ± 0.0345	1.149 ± 0.421	4.911 ± 0.675	−0.361 ± 0.454
0.03	−0.13	0.14	−2.944 ± 0.836	3.607 ± 0.760	1.065 ± 0.336	1.936 ± 0.593	4.855 ± 0.966	2.126 ± 0.334	2.341 ± 0.239
18626	2.041 ± 0.039		4.730 ± 0.713	3.115 ± 0.351	0.0600 ± 0.0117	0.1609 ± 0.0151	...	2.317 ± 0.380	1.687 ± 0.264
A2734	0.080		0.505 ± 0.239	...	...	...	...	...	...
NFP J001024.3−284935			0.272 ± 0.645	−1.753 ± 0.988	0.0553 ± 0.0267	0.0812 ± 0.0330	1.736 ± 0.448	4.635 ± 0.766	−1.225 ± 0.479
0.10	0.41	0.17	−5.047 ± 0.870	4.263 ± 0.895	1.193 ± 0.443	3.046 ± 0.622	4.245 ± 1.072	1.909 ± 0.363	2.030 ± 0.245
17923	2.234 ± 0.041		5.946 ± 0.764	3.599 ± 0.390	0.0563 ± 0.0119	0.2086 ± 0.0156	...	3.133 ± 0.471	1.768 ± 0.346
A2734	0.080		0.998 ± 0.279	...	...	...	...	...	...
NFP J001032.5−285154			0.309 ± 0.781	−0.877 ± 1.168	0.0538 ± 0.0326	0.0830 ± 0.0421	0.912 ± 0.502	5.587 ± 0.799	−1.787 ± 0.581
−0.17	−0.04	0.19	−6.479 ± 1.041	5.265 ± 1.049	1.230 ± 0.439	4.227 ± 0.723	6.701 ± 1.138	1.282 ± 0.383	1.531 ± 0.316
17441	2.156 ± 0.047		5.538 ± 0.898	4.494 ± 0.422	0.1003 ± 0.0129	0.2501 ± 0.0168	...	2.553 ± 0.476	1.312 ± 0.316
A2734	0.080		1.051 ± 0.284	...	...	...	...	...	...
NFP J001041.8−283444			0.032 ± 1.019	−2.041 ± 1.502	0.0132 ± 0.0413	0.0593 ± 0.0358	1.862 ± 0.589	4.676 ± 0.892	−2.127 ± 0.696
−0.08	0.01	0.21	−5.915 ± 1.202	2.918 ± 1.555	0.692 ± 0.527	3.183 ± 0.767	3.673 ± 1.397	2.344 ± 0.511	2.435 ± 0.349
18190	1.988 ± 0.053		4.396 ± 0.987	2.885 ± 0.527	0.0140 ± 0.0135	0.1603 ± 0.0179	...	2.059 ± 0.583	1.595 ± 0.376
A2734	0.080		0.436 ± 0.300	...	...	...	...	...	...
NFP J001042.6−284916			−0.028 ± 0.723	−1.420 ± 1.070	0.0459 ± 0.0282	0.0712 ± 0.0332	1.368 ± 0.443	4.639 ± 0.756	−0.680 ± 0.515
−0.60	−0.80	0.20	−3.697 ± 0.933	4.181 ± 0.960	0.949 ± 0.415	2.020 ± 0.796	6.930 ± 1.133	1.633 ± 0.415	1.843 ± 0.309
18682	1.942 ± 0.047		4.511 ± 0.903	3.647 ± 0.454	0.0510 ± 0.0129	0.2026 ± 0.0168	...	1.500 ± 0.472	1.817 ± 0.370
A2734	0.080		0.760 ± 0.287	...	...	...	...	...	...

Notes.—Table 5 is published in its entirety in the electronic edition of the *Astrophysical Journal*. A portion is shown here for guidance regarding content.

5.2. *Analysis of Individual Line Strengths with Velocity Dispersion*

We restrict our attention to the 20 line-strength indices that are least affected by systematic errors. In fitting our line-strength–σ data, data points more than 4 σ away from the initial weighted least-squares fit are excluded. We then perform a linear regression of line strength on $\log\sigma$ on the remaining data, allowing for measurement errors in both variables but assuming that any intrinsic scatter comes solely from the line strengths. Table 7 summarizes the parameters of our fits. Slopes and errors are quoted at both full (instrumental) and Lick resolution. In Figure 8, the line-strength–σ relations for 20 indices are shown, both for individual data points and for means in velocity dispersion bins. The solid lines show the fitted slopes.

Of all the line indices, magnesium correlates most strongly with σ and has been most widely used as an indicator of stellar populations. In order to compare our Mg-σ fit with those from other surveys, we convert our Mg b line in Å to Mg b' in mag using the conversion defined in Colless et al. (1999):

$$\text{Mg } b' = -2.5\log_{10}\left(1-\frac{\text{Mg } b}{32.5}\right). \quad (3)$$

For ~3400 galaxies, the Mg b' line strength is fitted with the relation $\Delta\text{Mg } b' \propto (0.122 \pm 0.002)\Delta(\log\sigma)$. This slope is very close to that found by EFAR (0.13 ± 0.017), SDSS (0.15 ± 0.02), and Kuntschner et al. (2001) (0.142 ± 0.013). When the data from Kuntschner et al. (2001) are fitted using our fitting method, the slope (0.113) is closer to ours. Our intrinsic scatter of 0.011 mag is lower than those of both EFAR and Kuntschner et al. (2001); we do not compare our scatter to that of SDSS since they effectively reduce their scatter by creating composite spectra from individual galaxies.

We take the logarithm of our Hβ line strengths to compare the subsequent trends with σ with other surveys: our Hβ fit yields $\log H\beta \propto (-0.306 \pm 0.008)\log\sigma$ for 3440 galaxies compared to a slope of −0.24 ± 0.03 for SDSS. Jorgensen (1997) found $\log H\beta \propto (-0.23 \pm 0.08)\log\sigma$ with an intrinsic scatter of 0.061 compared to our intrinsic scatter of 0.041. We discuss the effects on our conclusions of a different Hβ slope in § 6.3.

In order to compare our iron relations with other surveys, we defined $\langle\text{Fe}\rangle$ as

$$\langle\text{Fe}\rangle = \frac{\text{Fe5270} + \text{Fe5335}}{2}. \quad (4)$$

For 1090 galaxies with both Fe5270 and Fe5335 measurements, our best fit to $\log\sigma$ is $\log\langle\text{Fe}\rangle \propto (0.088 \pm 0.008)\log\sigma$. This is consistent with the lower value found by Jorgensen (1997) (0.075 ± 0.025) and the higher value in the SDSS (0.11 ± 0.03).

The line strength Fe5709 has a negative slope against $\log\sigma$ and is the only iron line to behave in this manner. While it is possible that the line would be contaminated by strong sky emission from the Na I $\lambda\lambda$5683, 5688 lines, this would only affect measurements in the extremely low redshift clusters. We have

TABLE 6
FULL-RESOLUTION LINE-STRENGTH INDICES AND ASSOCIATED DATA

Galaxy ID			$H\delta_F$	$H\delta_A$	CN_1	CN_2	Ca4227	G4300	$H\gamma_F$
W(O III)	$W(H\beta)$	$\varepsilon(W)$	$H\gamma_A$	Fe4383	Ca4455	Fe4531	Fe4668	$H\beta$	$H\beta^+$
$cz_\odot$	$\log\sigma$		Fe5015	Mg b	Mg_1	Mg_2	Fe5270	Fe5335	Fe5406
Cluster	$\log(d/d_0)$		Fe5709	Fe5782	Na5895	TiO_1	TiO_2	$H\alpha_A$	$H\alpha_F$
NFP J001006.3−284623			2.186 ± 0.977	−1.756 ± 1.561	0.0355 ± 0.0418	0.0824 ± 0.0493	0.141 ± 0.692	7.007 ± 1.039	−0.874 ± 0.703
−0.06	−0.35	0.22	−4.250 ± 1.221	5.697 ± 1.330	0.885 ± 0.755	3.795 ± 0.909	4.022 ± 1.343	2.189 ± 0.499	2.211 ± 0.338
18989	1.877 ± 0.050		4.972 ± 1.033	3.213 ± 0.475	0.0448 ± 0.0129	0.1671 ± 0.0164	...	3.212 ± 0.559	1.016 ± 0.434
A2734	0.080		1.333 ± 0.338	...	...	...	...	...	...
NFP J001010.7−285020			0.703 ± 0.689	−1.443 ± 1.079	0.0275 ± 0.0306	0.0515 ± 0.0366	1.345 ± 0.498	5.296 ± 0.783	−0.446 ± 0.491
0.03	−0.13	0.14	−3.011 ± 0.850	4.431 ± 0.941	1.551 ± 0.479	2.328 ± 0.701	5.154 ± 1.025	2.185 ± 0.361	2.322 ± 0.246
18626	2.041 ± 0.039		5.426 ± 0.822	3.179 ± 0.371	0.0609 ± 0.0110	0.1617 ± 0.0143	...	2.829 ± 0.456	2.094 ± 0.343
A2734	0.080		0.546 ± 0.274	...	...	...	...	...	...
NFP J001024.3−284935			0.243 ± 0.677	−2.122 ± 1.066	0.0744 ± 0.0312	0.1053 ± 0.0367	1.929 ± 0.507	4.841 ± 0.850	−1.280 ± 0.514
0.10	0.41	0.17	−4.579 ± 0.880	4.710 ± 1.004	1.509 ± 0.555	3.436 ± 0.706	4.078 ± 1.048	1.934 ± 0.385	2.055 ± 0.258
17923	2.234 ± 0.041		7.011 ± 0.902	3.674 ± 0.409	0.0589 ± 0.0114	0.2105 ± 0.0149	...	3.641 ± 0.543	1.911 ± 0.390
A2734	0.080		1.067 ± 0.310	...	...	...	...	...	...
NFP J001032.5−285154			0.490 ± 0.809	−1.176 ± 1.242	0.0668 ± 0.0372	0.0923 ± 0.0440	1.185 ± 0.646	6.150 ± 0.924	−1.862 ± 0.612
−0.17	−0.04	0.19	−6.273 ± 1.036	5.713 ± 1.152	1.811 ± 0.637	4.644 ± 0.800	6.825 ± 1.166	1.477 ± 0.457	1.438 ± 0.308
17441	2.156 ± 0.047		6.469 ± 1.051	4.668 ± 0.449	0.1047 ± 0.0125	0.2541 ± 0.0161	...	3.193 ± 0.589	1.727 ± 0.426
A2734	0.080		1.237 ± 0.344	...	...	...	...	...	...
NFP J001041.8−283444			0.079 ± 1.049	−2.570 ± 1.570	0.0279 ± 0.0415	0.0929 ± 0.0489	2.060 ± 0.657	5.373 ± 1.063	−2.085 ± 0.722
−0.08	0.01	0.21	−5.704 ± 1.222	2.755 ± 1.470	0.996 ± 0.747	4.164 ± 0.983	3.711 ± 1.406	2.386 ± 0.532	2.425 ± 0.356
18190	1.988 ± 0.053		5.083 ± 1.145	2.890 ± 0.536	0.0154 ± 0.0136	0.1604 ± 0.0173	...	2.252 ± 0.637	1.890 ± 0.458
A2734	0.080		0.539 ± 0.383	...	...	...	...	...	...
NFP J001042.6−284916			0.077 ± 0.758	−1.731 ± 1.139	0.0586 ± 0.0329	0.0961 ± 0.0382	1.571 ± 0.514	4.958 ± 0.858	−0.832 ± 0.546
−0.60	−0.80	0.20	−3.647 ± 0.947	4.662 ± 1.085	1.342 ± 0.577	2.329 ± 0.890	6.974 ± 1.150	1.751 ± 0.460	1.754 ± 0.304
18682	1.942 ± 0.047		4.843 ± 0.982	3.684 ± 0.469	0.0529 ± 0.0125	0.2069 ± 0.0163	...	1.864 ± 0.580	2.101 ± 0.441
A2734	0.080		0.869 ± 0.340	...	...	...	...	...	...

NOTES.—Table 6 is published in its entirety in the electronic edition of the *Astrophysical Journal*. A portion is shown here for guidance regarding content.

analyzed the line-strength data from Trager et al. (1998) and, for Fe5709, find a slope of -0.06 ± 0.09, which is consistent with our result.

We have shown that each of the 20 line strengths considered shows significant dependence on velocity dispersion. The pattern that emerges is that all Balmer lines decrease with increasing velocity dispersion, whereas all other lines (with the exception of Fe5709) increase with increasing velocity dispersion. The following section describes a robust method for interpreting these trends as scaling relations of stellar population properties, as a function of galaxy mass.

6. GLOBAL AGE AND METALLICITY TRENDS

Stellar population models, such as those by TMB03 and the extension of these that includes $H\gamma$ and $H\delta$ (Thomas et al. 2004), predict line strengths from a range of stellar population parameters. These models are an extension of those by Maraston (1998) with adjustments based on theoretical stellar atmosphere calculations. TMB03 currently provide the only models that predict line strengths for populations with nonsolar abundance ratios, as required for elliptical galaxy studies. TMB03 do not, however, include possible contributions from blue horizontal branch (BHB) stars, which might arise from a low-metallicity subpopulation (Maraston & Thomas 2000) or from enhanced mass loss in evolved stars (Thomas et al. 2005). Whatever their origin, such stars would contribute to a strengthening of the Balmer absorption lines, mimicking the effect of younger ages. In principle, BHB stars have greater impact on the high-order lines than on $H\beta$, which could ultimately provide a means to distinguish their effects from those of a younger population (Schiavon et al. 2004). In this paper we assume that the BHB contribution is either negligible or at least not strongly dependent on galaxy mass.

6.1. *Method*

We wish to use our line-strength data to determine ages, metallicities, and α-element enhancements as a function of galaxy mass. There are several ways to extract this information. The usual method, adopted by most previous studies, is to derive these parameters for individual galaxies by interpolating (sometimes extrapolating) the model grids. However, when the measurement errors in the line strengths are nonnegligible (as is the case with the NFPS data), the tilt of the model grid leads to correlated errors in age and metallicity for each galaxy (Kuntschner et al. 2001). These correlated errors complicate the interpretation of the data, potentially generating a spurious age-metallicity correlation (Terlevich & Forbes 2002).

This problem can be overcome by stacking the spectra of similar galaxies, for example, galaxies within a narrow bin of velocity dispersion (Bernardi et al. 2003), to create a composite spectrum of high S/N. Equivalently, one can average the line-strength measurements themselves for galaxies in each bin. In either case, one averages over scatter in the line strengths that may be due to measurement errors but also averages the scatter

TABLE 7
LINE-STRENGTH–σ FITS

Line Strength	Units	N_{gal}[a]	Mean Slope (Lick)[a]	Mean Slope (Inst)[b]
$H\delta_A$	Å	3419	-3.484 ± 0.109	-3.381 ± 0.116
$H\delta_F$	Å	3399	-1.444 ± 0.055	-1.408 ± 0.065
CN_1	mag	3435	0.197 ± 0.003	0.193 ± 0.004
Ca4227	Å	3420	0.912 ± 0.039	0.421 ± 0.042
$H\gamma_A$	Å	3428	-3.427 ± 0.096	-2.997 ± 0.099
$H\gamma_F$	Å	3434	-2.163 ± 0.052	-2.222 ± 0.056
Fe4383	Å	3437	1.448 ± 0.084	0.990 ± 0.088
Ca4455	Å	3457	0.631 ± 0.033	0.451 ± 0.040
Fe4531	Å	3442	0.917 ± 0.048	0.643 ± 0.056
Fe4668	Å	3443	5.230 ± 0.109	4.885 ± 0.112
$H\beta$	Å	3450	-1.171 ± 0.032	-1.166 ± 0.037
Fe5015	Å	3450	1.029 ± 0.068	1.008 ± 0.077
Mg_1	mag	2018	0.121 ± 0.003	0.123 ± 0.002
Mg_2	mag	1993	0.189 ± 0.003	0.184 ± 0.003
Mg b	Å	3414	3.201 ± 0.041	3.047 ± 0.043
Fe5270	Å	1954	0.620 ± 0.041	0.614 ± 0.045
Fe5335	Å	2113	0.821 ± 0.045	0.854 ± 0.053
Fe5406	Å	2825	0.432 ± 0.029	0.432 ± 0.033
Fe5709	Å	2503	-0.131 ± 0.022	-0.096 ± 0.026
Fe5782	Å	2382	0.172 ± 0.024	0.289 ± 0.027
Na5895	Å	1514	4.276 ± 0.075	4.590 ± 0.084
TiO_1	mag	789	0.021 ± 0.002	0.021 ± 0.002
TiO_2	mag	916	0.046 ± 0.002	0.046 ± 0.003
$H\alpha_A$	Å	542	-0.722 ± 0.148	-0.791 ± 0.067
$H\alpha_F$	Å	684	-0.952 ± 0.104	-1.147 ± 0.099

[a] At Lick resolution.
[b] At full (instrumental) resolution.

due to intrinsic (possibly correlated) dispersion in the galaxy population parameters. The obvious weakness of this approach is that one can only state confidently the characteristics of the mean galaxy; thus, there is no information regarding the variation in parameters from galaxy to galaxy within the bin.

As a preliminary step, we plot our binned line strengths on top of grids derived from line-strength–parameter scaling relations from TMB03 in Figure 9. The quantity $[\mathrm{MgFe}]'$ in the bottom two panels was defined by TMB03 as

$$[\mathrm{MgFe}]' \equiv \sqrt{\mathrm{Mg}\ b(0.72 \times \mathrm{Fe5270} + 0.28 \times \mathrm{Fe5335})}. \quad (5)$$

This index was defined to be sensitive to overall metallicity but almost independent of $[\alpha/\mathrm{Fe}]$. In Figure 9, the area inside the dotted trapezoid represents the average error for a *single* line-strength measurement in each bin. Since each bin represents $\sim$700 galaxies, the errors in the mean values are very small and not plotted here.

In the top left panel of Figure 9, where we plot models with $[\alpha/\mathrm{Fe}] = 0.0$, our five points form a line of almost constant metallicity $[Z/\mathrm{H}] = 0.0$ along with an increasing age estimate as the velocity dispersion increases. In the top right panel where $[\alpha/\mathrm{Fe}] = 0.3$, we can see a slight shift of our data points toward higher metallicity and again toward older ages, with increasing σ. In the bottom panels, the grids show predictions for constant age, while our five binned points follow a rough trend in metallicity: increasing $[Z/\mathrm{H}]$ with increasing velocity dispersion. There is a slightly less pronounced trend of increasing $[\alpha/\mathrm{Fe}]$ with increasing σ. It should be realized that, since each panel shows a two-parameter projection of the three-parameter models, it is not trivial to read off the "correct" values from these plots.

The stacking methods described above work well for a triplet of line strengths (e.g., Hβ, Mg b, $\langle$Fe$\rangle$), from which one derives the triplet of galaxy parameters age, metallicity, and α-enhancement. However, in general, each choice of line-strength triplet will lead to different stellar population parameters. Ideally, we seek a method to combine the information from all line strengths simultaneously. An obvious approach is to fit, via a χ^2 method, all line strengths as a function of the galaxy parameters (Proctor et al. 2004a). Unfortunately, the zero points are typically uncertain, both for the model predictions and for the observational measurements. As a result, the model grids often systematically under- or overpredict certain line strengths. Inverting the grids then yields estimated ages (and other parameters) that are systematically in error. In a χ^2 approach, the line strengths that are affected in this way will be outliers in the fit and will carry disproportionate weight in the results.

Here we choose instead to describe the whole sample simultaneously, in terms of a set of scaling relations, with log (age), $[Z/\mathrm{H}]$, and $[\alpha/\mathrm{Fe}]$ each following a linear trend with $\log\sigma$. Here $[Z/\mathrm{H}]$ is the total metallicity, expressed logarithmically relative to solar, while $[\alpha/\mathrm{Fe}]$ is the α-element enhancement, similarly expressed. These global trends can be estimated by modeling their combined effects on the observed line-strength–σ relation. We can find the parameter-σ relations from

$$\frac{dI}{d\log\sigma} = \sum_{i=1}^{N} \frac{dI}{dP}\frac{dP}{d\log\sigma}, \quad (6)$$

where I is the line strength (at Lick resolution), P is a parameter from the stellar population models (log (age), $[Z/\mathrm{H}]$, or $[\alpha/\mathrm{Fe}]$), $dI/d\log\sigma$ are our line-strength–σ slopes from the previous section, and dI/dP are the line-strength–parameter "responses" from the TMB03 models. While the absolute zero point of the grids is uncertain, the dI/dP responses are more robust. We emphasize that our $dI/d\sigma$ relations involve the averaging of $\sim$3000 line-strength measurements for each index; thus, statistical errors are suppressed and do not significantly influence the results of the regression. The aim of this analysis is to constrain the $dP/d\log\sigma$ representing the slope of the parameter scaling relations.

This differential method has a number of advantages over the "grid inversion" approach. As already mentioned, it avoids the problem of interpreting correlated errors in the derived parameters. Also, it uses only the *relative* changes in the predicted line strengths from the TMB03 models and the *relative* changes in our measured line strengths with velocity dispersion. Our method is thus explicitly insensitive to calibration uncertainties in the models and to overall zero-point errors in the line-strength measurements. Finally, unlike stacking spectra or line-strength measurements by velocity dispersion, there is no need to bin the data into arbitrarily defined subsets. The main shortcomings of the method are its assumptions that the model grids are parallel over the parameter space spanned by our sample and that the parameter–$\log\sigma$ relations are indeed linear.

The outputs of the differential method are the parameter-σ scaling relations, which depend on the choice of the input model responses dI/dP. To estimate dI/dP, we calculate $\Delta I/\Delta P$ from the model grids themselves by differencing an upper (or lower) value with a fiducial value. For example, our data span 3–15 Gyr in age, with a central age of 8 Gyr. Referring to Figure 9, for a metallicity of $[Z/\mathrm{H}] = 0.0$ and $[\alpha/\mathrm{Fe}] = 0.0$, we see that $\Delta\mathrm{H}\beta_{\mathrm{pred}} = -0.288$ between 5 and 10 Gyr. Note that the grids are very nearly parallel. Thus, if we had assumed $[Z/\mathrm{H}] = 0.35$ instead, we would

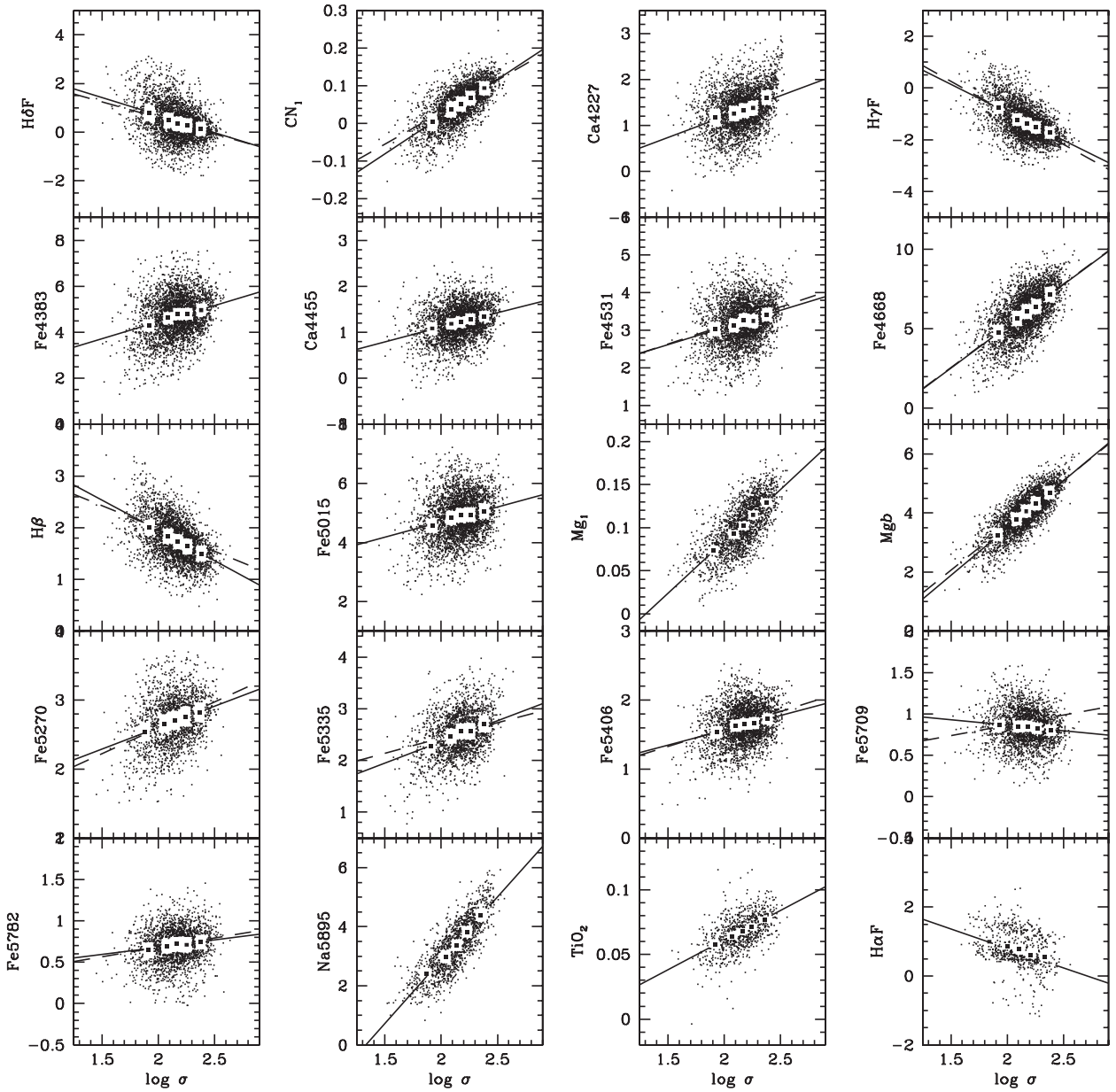

FIG. 8.—Our line-strength–σ relations for each index. The line strength and velocity dispersion for the galaxies (*points*) are sorted into five bins by increasing velocity dispersion (*black boxes outlined in white*). Solid lines are the slopes from the linear regression of line strength on $\log\sigma$. Dashed lines represent slopes predicted from our derived trends of [α/Fe], [Z/H], and age with velocity dispersion in § 6 (see text in that section).

have obtained a very similar value, $\Delta H\beta_{\rm pred} = -0.294$. To obtain the best estimates of the model responses, we need to make reasonable choices for the central values of the three parameters, as well as for their upper and lower values. We adopt as a central value the parameters from our fits corresponding to the median velocity dispersion ($\log\sigma \sim 2.1$): [α/Fe] = 0.2, age = 8 Gyr, and [Z/H] the average of 0.0 and 0.35. The range for each parameter was restricted to the range spanned by the sample galaxies, i.e., age of 3–15 Gyr, [α/Fe] of 0.0–0.5, and [Z/H] of -0.33 to 0.67. Because the grids are not exactly parallel, there is a small uncertainty in our derived parameter scaling relations introduced by our choice of central values. We discuss this in more detail in § 6.3.

To fit three model parameters, we need at least three line-strength–σ relations. We can overconstrain the fit by including more than three indices, and in practice we use the following 12 line indices: CN_1, Hβ, Hγ_F, Hδ_F, Fe4531, Fe4668, Mg b, Fe5270, Fe5335, Fe5406, Fe5709, and Fe5782. This set includes three age-sensitive Balmer lines, two indices quite sensitive to α-enhancement (Mg b, CN_1), and a range of mainly metallicity-sensitive features. We did not include lines such as Fe4383 and Fe5015 where the systematic correction to the line strength is a significant factor of the random error (see § 2.4). Ca4227 was excluded because in general the Ca abundances are not very well understood (Saglia et al. 2002), and we avoid including partially redundant indices, for instance, Hγ_A. Alternative choices of indices are possible, and the selection of which to include can affect the fit results by more than the formal random errors. We discuss this issue further in § 6.3.

The $dI/d\log\sigma$ slopes we use from our data derive from the regression of line strength on $\log\sigma$, as in Table 7. To weight contributions from the different line indices in the fit, we use the

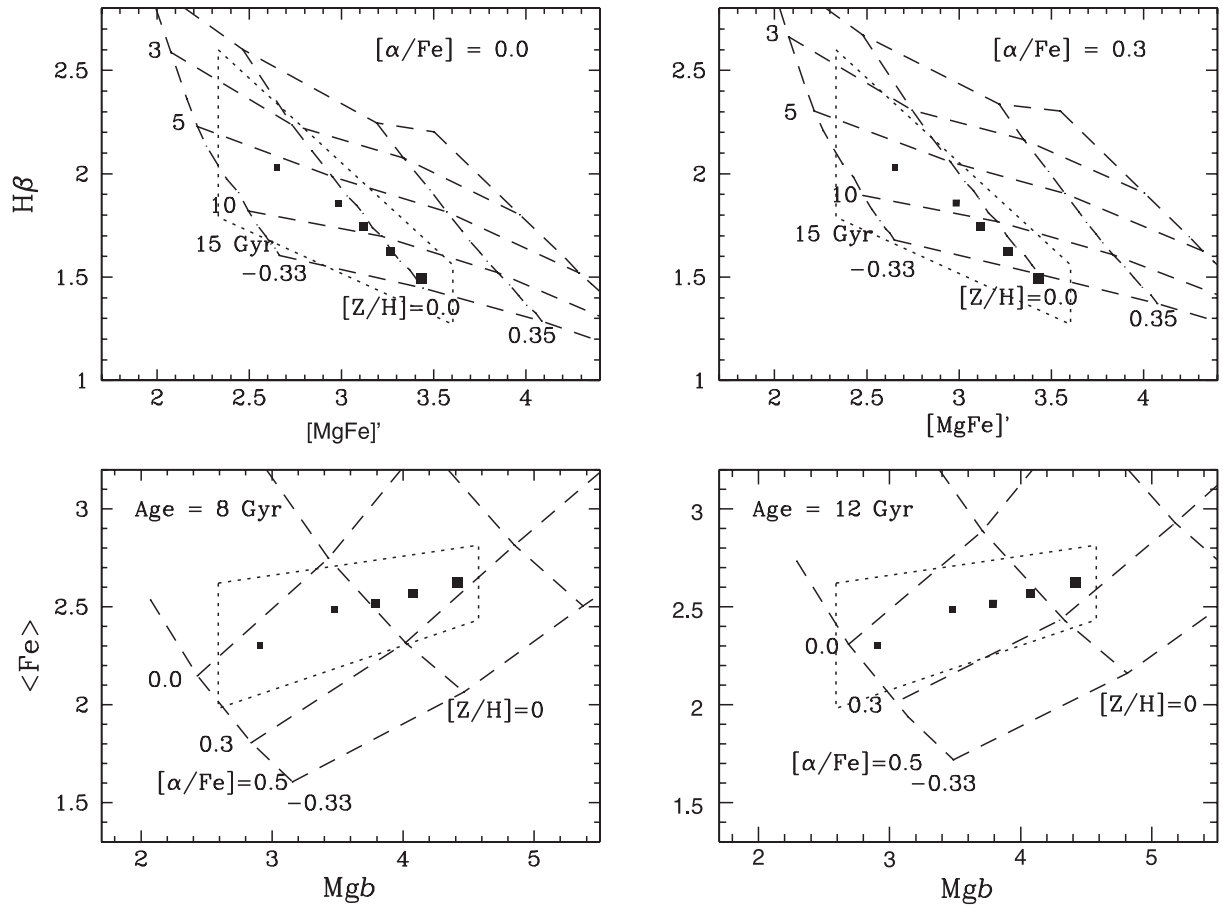

FIG. 9.—Model grids from stellar population models by TMB03. Each data point is the mean of a range of velocity dispersion, with each bin having approximately the same number of galaxies. The size of the points increases as σ increases. The area inside the dotted line is the range covered by the average errors of the line strengths. Line strengths are in units of Å.

formal regression errors on the line-strength–σ slopes. No contribution from model error is included.

6.2. *Results*

When all 12 line indices are used in the fit, we obtain $\alpha/\mathrm{Fe} \propto \sigma^{0.31}$, $Z/\mathrm{H} \propto \sigma^{0.53}$, and age $\propto \sigma^{0.59}$. The formal error estimates are very small (0.01, 0.02, and 0.02, respectively) and, as we discuss in greater detail below, are not representative of the true uncertainties. For more insight into which line strengths drive the fit results, we have explored restricted models in which one or more of the parameters are held constant with σ. In Figure 10 we show the predicted versus observed line-strength–σ slopes for the 12 line indices when different combinations of the stellar parameters are allowed to vary with σ. We normalize both the "actual slope" (i.e., our measured line-strength–σ slope) and the predicted slope by dividing by the error in our line-strength–σ slopes. Thus, the diagonal line in each panel shows where the actual and predicted slopes agree. Any vertical displacement of a line strength indicates an under- or overprediction of the observed slope.

Figure 10*a* shows good agreement between the actual and predicted line-strength–σ slopes when [α/Fe], [Z/H], and age all vary with σ, i.e., our default solution. When age is forced to be constant with σ, as in Figure 10*b*, the predicted (negative) slopes of the Balmer lines are too shallow and the predicted (positive) slopes of the α-sensitive indices, Mg b and CN_1, are also too low. Thus, without an age-σ relation, we cannot simultaneously reproduce both the Balmer lines and the α-sensitive slopes.

We also see that if metallicity is held constant with σ, as in Figure 10*d*, the predicted slopes of the Balmer lines are too steep, as is that of Mg b. Similar arguments apply to the other restricted fits. In the case where only [α/Fe] is allowed to vary with σ (Fig. 10*g*), the fit fails catastrophically, since a correctly predicted Mg b slope would require [α/Fe] increasing with σ, while a decreasing [α/Fe] with σ would be needed to produce the observed negative Balmer line slopes. The result is a very poor fit with no [α/Fe] trend at all.

Using the derived parameter-σ relations, we can rederive our line-strength–σ fits and check the consistency with our data-derived line-strength–σ slopes. We used our default solution for the parameter-σ relations ($\alpha/\mathrm{Fe} \propto \sigma^{0.31}$, $Z/\mathrm{H} \propto \sigma^{0.53}$, and age $\propto \sigma^{0.59}$) to generate the predicted slopes (*dashed lines*) in Figure 8. For each line strength, the predicted slope is very similar to the measured slope with the exception of Fe5709, whose measured line-strength–σ slope is opposite to those of the other iron line indices.

In § 5 we noted that the Hβ slopes found by both SDSS and Jorgensen (1997) are 20% lower than ours. We have repeated the analysis using their slope but found no significant change in our results, including the age-mass gradient.

6.3. *Error Estimation*

The formal errors on the galaxy parameter slopes $dP/d(\log\sigma)$ are very small: specifically, we find $\alpha/\mathrm{Fe} \propto \sigma^{0.31\pm0.01}$, $Z/\mathrm{H} \propto \sigma^{0.53\pm0.02}$, and age $\propto \sigma^{0.59\pm0.02}$. From a purely statistical point of view, these errors may be slightly underestimated for two

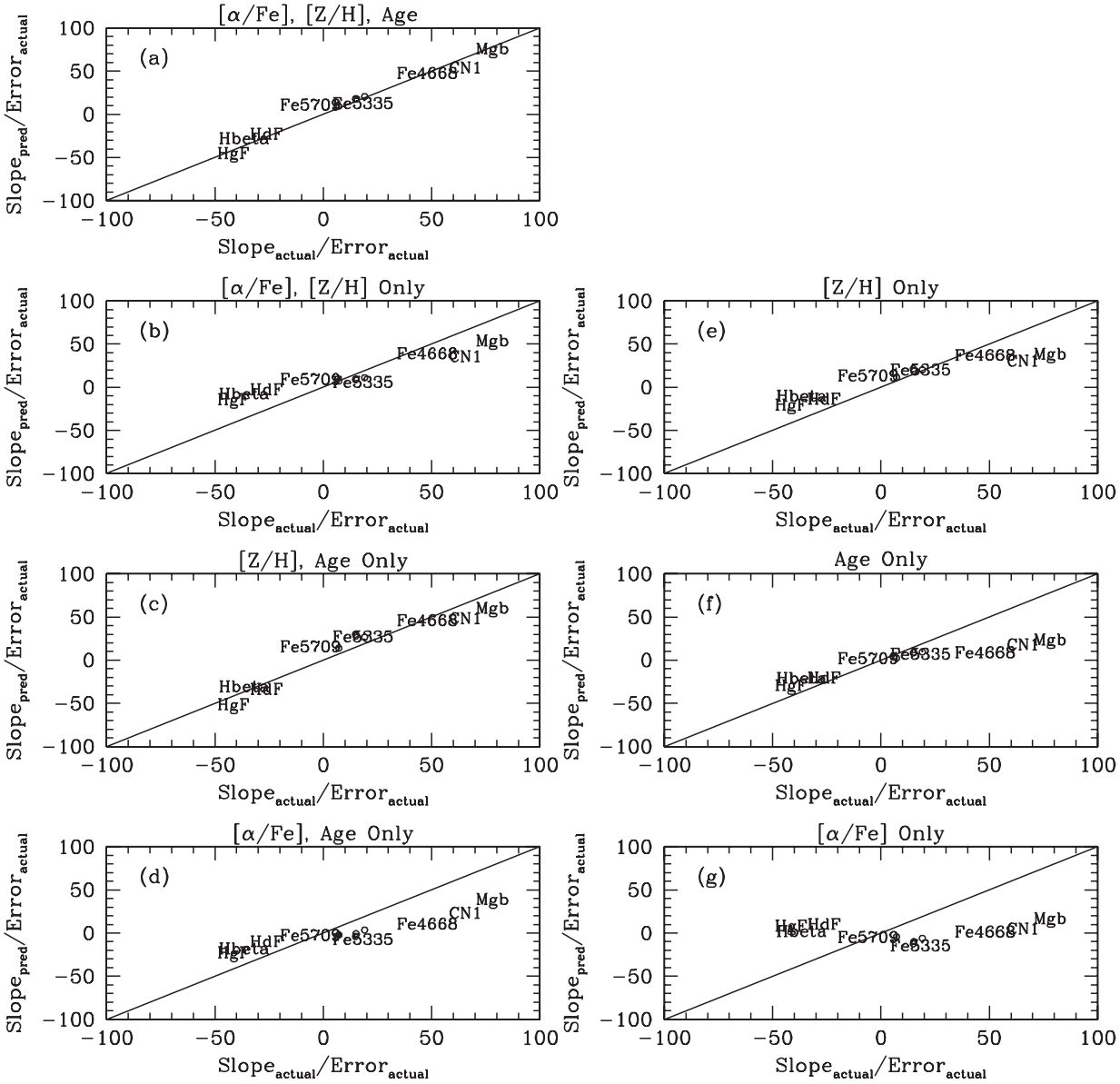

FIG. 10.—Comparison of the predicted index-σ slopes (*vertical axis*) with the observed slopes (*horizontal axis*) for different combinations of the three stellar population parameters. Labels indicate the results for each of the 12 line-strength indices. The diagonal line indicates perfect agreement between model-predicted and measured slopes. Note that both model-predicted slopes and measured slopes are normalized by the formal error in the measured slope. Thus, the vertical offset from the line indicates the degree to which the model fails to predict the observed slope. Panel (*a*) shows our default solution in which [α/Fe], [*Z*/H], and log (age) are scaled with log σ. The model predictions for Fe5709 deviate from the observed slope at a high significance level (compare with Fig. 8). Note, however, that Fe5709 is discrepant in all panels. Panels (*b*)–(*d*) show the comparisons when we assume no trend with σ for age, [α/Fe], and [*Z*/H], respectively. In these cases, there are multiple indices that deviate strongly from the model predictions. For example, if we assume no age gradient (panel [*b*]), we cannot simultaneously reproduce the slopes of Mg *b* and CN_1 and the Balmer lines. The fits are considerably poorer when only one parameter is allowed to vary, as in panels (*e*)–(*g*).

reasons: first, the line-strength measurements are not strictly independent since in a few cases the bandpasses overlap (e.g., Fe5270 and Fe5335); and second, it is possible that the scatter in intrinsic parameters from galaxy to galaxy may be correlated (e.g., Trager et al. [2000b] suggest that at a given σ, younger galaxies are also more metal-rich). Neither of these effects will bias the resulting age, metallicity, and α-enhancement trends, which are essentially the means of the galaxy parameters at a fixed σ. However, we show below that systematic effects vastly dominate over the statistical errors, so we use the former as our error estimate and neglect the latter. There are two sources of systematic error: one is the choice of line-strength indices used in the regression, and the other is due to the uncertainty in estimating the responses *dI*/*dP*.

To estimate the systematic errors due to choice of line-strength indices, we divide our 12 line indices into three groups: Balmer lines (Hβ, Hγ_F, Hδ_F), α-sensitive indices (Mg *b*, CN_1), and the seven Fe-dominated indices. We then examine the scaling relations that result when only one index from a given group is used and all indices from the other two groups are used. Using Figure 11 as a guide, we can draw several conclusions. First, the only consistently "outlying" line indices are Fe5335 and Fe5782, which prefer a weaker age trend, a stronger metallicity trend, and a weaker [α/Fe] trend, and Fe5709, for which the reverse is true. In general, there is little spread of the scaling relations; choosing other combinations of line-strength indices yields consistent results. In particular, note that the age gradient is essentially unchanged whichever Balmer line is used. Note

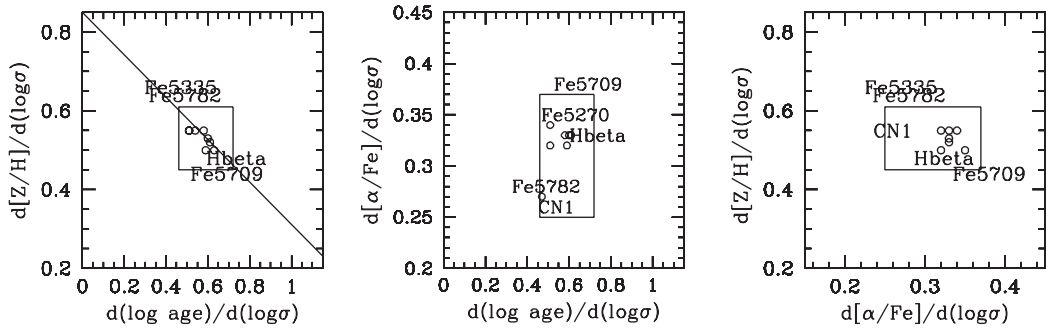

FIG. 11.—Effect of inclusion of certain line strengths on our regression. Each label or open circle represents the inclusion of that element and the exclusion of the other elements *in that group* (see text). For example, the Fe5709 label indicates the fit results when Fe5709 is the only Fe-dominated index included (all of the Balmer and α-sensitive indices are retained for this test, however). Similarly, the CN_1 label indicates that for that fit, CN_1 was the *only* α-index included (i.e., Mg *b* was excluded), but all of the Balmer and Fe indices were also included. For clarity, only indices that have a significant effect on the default solution are labeled; other indices are shown by open circles. The straight line in the left panel is drawn by eye to represent the linearity of the age-metallicity degeneracy. The rectangle in each panel outlines our default solution with error range.

also that the scatter along the age-metallicity degeneracy line is such that if the age trend is steeper by 0.1, the metallicity trend is flatter by $\sim$0.03. We estimate the resulting errors on our scaling relations by calculating the standard deviation of the results from all of the line-strength combinations. We find the deviations to be 0.03 in the $[\alpha/\mathrm{Fe}]$–$\log\sigma$ relation, 0.06 in $[Z/\mathrm{H}]$–$\log\sigma$, and 0.08 in age–$\log\sigma$.

Our second error component is the deviation of our parameter-σ scaling relations when different sets of *dI*/*dP* values were used in the regression. We consider eight cases in which our central values of $[\alpha/\mathrm{Fe}]$, $[Z/\mathrm{H}]$, and age are perturbed by $^{+0.2}_{-0.1}$, ±0.35, and ±2 Gyr, respectively, and the responses *dI*/*dP* accordingly recalculated around these central values. We find a standard deviation of 0.07 in the $[\alpha/\mathrm{Fe}]$–$\log\sigma$ slope, 0.08 in $[Z/\mathrm{H}]$–$\log\sigma$, and 0.10 in log (age)–$\log\sigma$. When these errors are added in quadrature with the errors from different line-strength combinations, the total error is 0.06 in $[\alpha/\mathrm{Fe}]$–$\log\sigma$, 0.08 in $[Z/\mathrm{H}]$–$\log\sigma$, and 0.13 in log (age)–$\log\sigma$.

In summary, we see that the inclusion or exclusion of individual indices and the choice of input *dI*/*dP* values affect the results by considerably more than the formal errors. The scatter between various results allows estimation of more realistic uncertainties. Our final results, with errors, are $\alpha/\mathrm{Fe} \propto \sigma^{0.31\pm0.06}$, $Z/\mathrm{H} \propto \sigma^{0.53\pm0.08}$, and age $\propto \sigma^{0.59\pm0.13}$. These errors are correlated in the sense that a stronger age trend corresponds to a weaker metallicity trend and vice versa. Figure 12 shows comparisons between our results from different combinations of the line indices and those from other studies; this figure is discussed further in § 7.2.

6.4. *Grid Inversion: A Consistency Check*

The slope analysis of the previous section makes the assumption that the grids are linear and parallel (i.e., that a single set of responses *dI*/*dP* is sufficient to capture the structure of the models). Also, we imposed a linear relationship of each population parameter with $\log\sigma$. We have argued that despite these restrictions, this explicitly differential method is more robust against calibration uncertainties in both the data and the models than direct interpolation of model grids and has the further advantage that it does not require explicit binning.

In this section we return to the grid interpolation method as a consistency test and to assess the limitations of the slopes analysis. Since the individual measurements are subject to sizable random errors, we use the line strengths averaged over the five velocity dispersion bins of Figure 9. The model grid is inverted by determining, for each velocity dispersion bin, the parameters [log (age), $[Z/\mathrm{H}]$, $[\alpha/\mathrm{Fe}]$] that best reproduce (according to a χ^2 statistic) the same 12 line indices used above. This analysis yields an estimate of the three population parameters as a function of $\log\sigma$ that is not forced to the linear form imposed in the slopes method. The results are shown in Figure 13. The grid inversion approach yields age-σ, $[Z/\mathrm{H}]$-σ, and $[\alpha/\mathrm{Fe}]$-σ relations in excellent agreement with results from our slopes method. At face value, the age of the most massive galaxies is $\sim$11 Gyr, but for the reasons emphasized elsewhere, the absolute values of age, $[Z/\mathrm{H}]$, and $[\alpha/\mathrm{Fe}]$ are less secure than the relative change along the mass sequence. The top panel of Figure 13 suggests a nonlinearity in the age–$\log\sigma$ relation, which appears to steepen

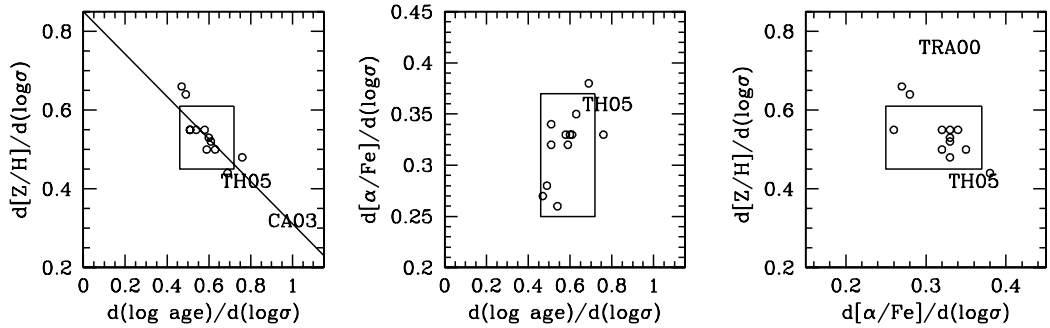

FIG. 12.—Our range of stellar population parameter scaling relations with σ plotted along with those from other groups (CA03 = our estimate from Fig. 21 and Table 9 from Caldwell et al. 2003; TRA00 = Trager et al. 2000b; TH05 = Thomas et al. 2005). Open circles represent the range in NFPS scaling relations when certain line strengths are included from the fit while the rest in their group have been excluded (see text and Fig. 11).

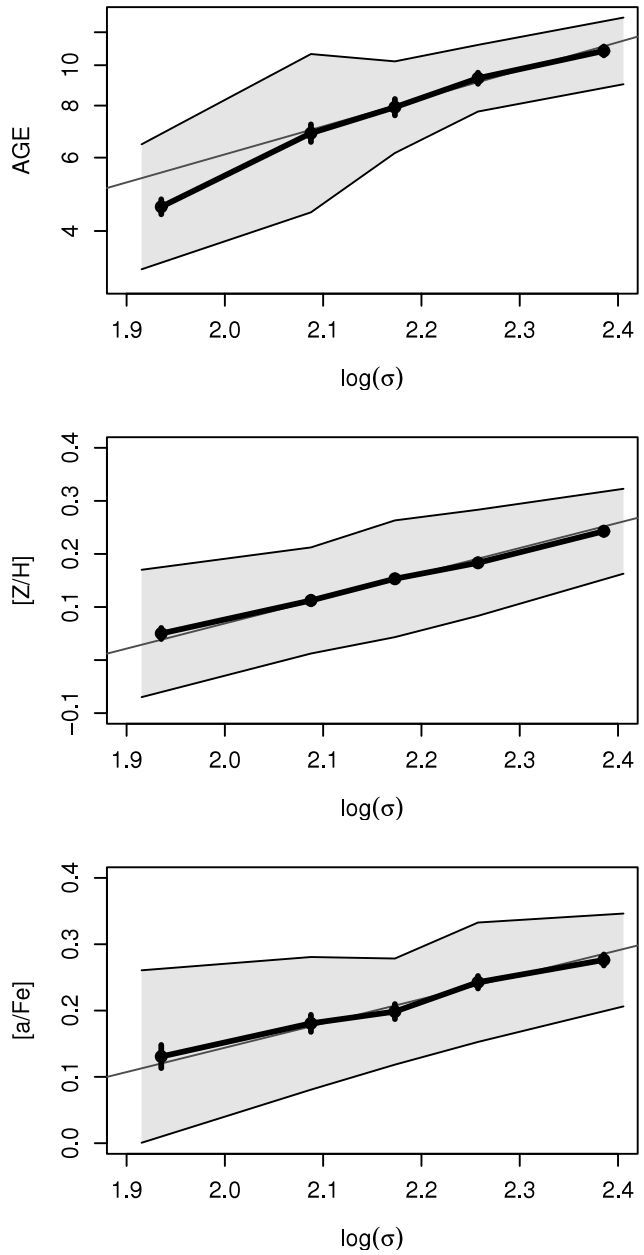

FIG. 13.—Fundamental scaling relations of the population parameters with σ. The thick line shows average age, metallicity, and [α/Fe] determined in bins of velocity dispersion, using the grid inversion method. The set of line indices employed is the same as that in the slopes method. For comparison, the linear relations obtained from the slopes method are overplotted as the thin line, with zero point adjusted to match the grid inversion results. The shaded regions indicate the *maximum* internal scatter estimated in § 6.5.

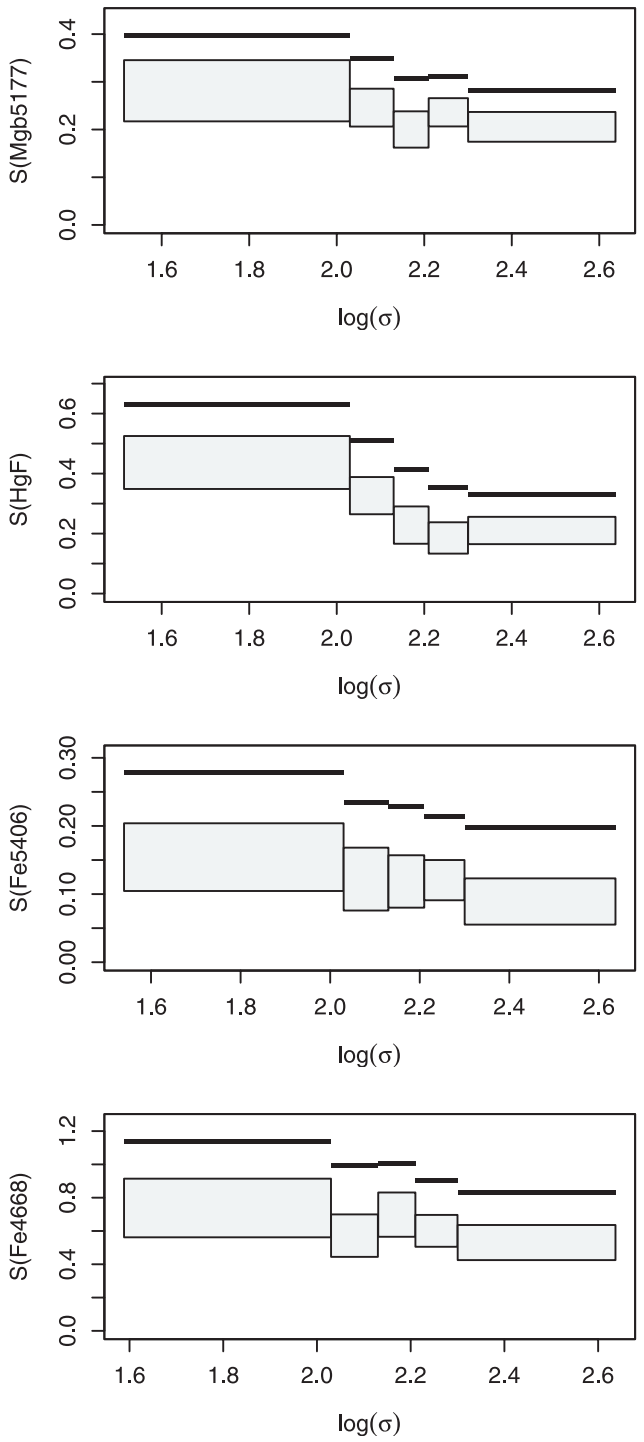

FIG. 14.—Total, measurement, and intrinsic scatter around the line-strength–σ relationships. In each panel, the black lines show the observed (i.e., total) scatter in Å (strictly, this is a robust estimate of the 1 σ scatter, based on the measured interquartile range). The 95% confidence intervals for the intrinsic 1 σ scatter are indicated by the shaded boxes.

by almost a factor of 2 at the lowest masses (i.e., low σ). Such behavior is, by definition, not observed in the slopes analysis of the previous sections.

6.5. *Internal Population Scatter*

An issue of considerable interest is the degree of internal scatter among galaxy properties at a given point on the mass sequence. For instance, the tightness of the color-magnitude relation has been used to infer limits on the spread in formation ages of elliptical galaxies (Bower et al. 1992). The scatter around line-strength–σ relations in principle provides very powerful constraints, since each line index has different sensitivities to the underlying population parameters.

While an upper limit to the scatter in line-strength–σ relations is readily available from the measured scatter, estimating the *intrinsic* scatter requires accurate knowledge of the experimental errors and a robust estimation method. In each bin, we

TABLE 8
AGE, METALLICITY, AND [α/Fe] BY σ BIN

Range of σ	$\langle \log \sigma \rangle$	$\langle$age$\rangle$	$\langle$[Z/H]$\rangle$	$\langle$[α/Fe]$\rangle$	$S(\log(\text{age}))$	S([Z/H])	S([α/Fe])
28–107	1.935	4.57	0.05	0.13	0.05–0.15	0.07–0.12	$\leq$0.13
107–135	2.088	6.87	0.11	0.18	0.08–0.19	0.06–0.10	$\leq$0.10
135–162	2.173	7.93	0.15	0.20	$\leq$0.11	0.07–0.11	$\leq$0.08
162–200	2.257	9.31	0.18	0.24	$\leq$0.08	0.06–0.10	$\leq$0.09
200–434	2.386	10.82	0.24	0.28	0.03–0.08	0.04–0.08	$\leq$0.07

first fit a linear line-strength–σ trend (allowing the slope and zero point to vary from bin to bin if the data require this). Then we model the line-strength residuals assuming a constant intrinsic scatter for that bin $S(I)$, in addition to the measurement errors (which include the systematic components estimated in § 2.4). In this way we can determine the range of $S(I)$ values that can reproduce the observed interquartile range of the data. Relative to a standard maximum likelihood estimate, this is more robust against outliers and focuses on matching the core of the distribution.

The 95% confidence limits on $S(I)$ are plotted in Figure 14 for a subset of our line indices, for the five velocity dispersion bins used elsewhere in this paper. The line indices shown are those in which intrinsic scatter is most cleanly detected: Mg b, H$\gamma_{\rm F}$, Fe5406, and Fe4668.

Given a set of models such as TMB03, the scatter $S(I)$ in each line index yields an upper and lower bound on the internal age scatter $S(\log(\text{age}))$, and similarly for [Z/H] and [α/Fe]. Using multiple line indices, these constraints can be compounded, to leave a narrow range in each parameter that is consistent with the observed scatter. In this way, we have determined the internal dispersion consistent with the observed scatters shown in Figure 14. To simplify the problem, we have assumed that the parameter scatters are independent, i.e., within each bin there is no internal correlation between age-σ and metallicity-σ residuals. In converting parameter scatter to predicted line-strength scatters, we use model responses appropriate to each bin, given the grid inversion results above. Note that we do not use Hβ here; significant intrinsic scatter in Hβ is observed, at a level inconsistent with the other line indices under any model for the underlying distribution. This is likely due to variations in low-level residual nebular emission.

The results of these calculations are included in Table 8, along with the results of the grid inversion analysis. For each bin, we tabulate the range and average of σ, the average stellar population parameters, and the range of population scatter consistent with the limits in Figure 14. The parameter scatters should be used with caution, since they are quite sensitive to the accuracy of our observational error estimates. Taken at face value, however, the analysis suggests a $\sim$25% scatter in metallicity at given σ, while the age scatter appears to vary with mass, increasing from $\sim$15% at high σ to $\sim$40% for the low-σ bin.

In the next section we discuss our results from this section in comparison with those of other studies and consider the relation of our results with recent observations of intermediate-redshift clusters.

7. DISCUSSION

7.1. *Interpretation of the Scaling Relations*

In the following sections we compare our results to other similar studies and to observations at higher redshift. First, however, it is useful to reiterate some caveats to interpreting the scaling relations derived here.

As in any survey, the sample selection criteria must be borne in mind when discussing our results. The NFPS sample analyzed here includes only cluster members, based on the criteria of Paper I. Within each cluster, the galaxies were selected by apparent magnitude (to $R = 17$) and color. Compared to some previous surveys, our sample probes to fairly low mass, $\sigma \sim 50$ km s^{-1}. The nonlinearity suggested by Figure 13 is such that samples of more massive objects would yield shallower age trends than studies covering the low-σ regime. The color criterion rejects blue galaxies farther than 0.2 mag from a red-sequence fit to each cluster. While such a cut should exclude actively star-forming galaxies, there is no explicit selection on morphological type. Thus, our sample includes many S0 galaxies and some bulge-dominated spiral galaxies; such galaxies were likely excluded (on a subjective basis) from many studies of "bona fide" elliptical galaxies.

To test for differences in the scaling relations between morphological classes, we have used GIM2D (Simard et al. 2002) to derive bulge-to-total light ratios for about half of the galaxies in our sample. We see indications that "diskier" galaxies (i.e., those with a bulge-to-total ratio $B/T \lesssim 0.5$) follow a steeper age-σ relation. However, there remains a significant age-σ trend for the bulge-dominated galaxies. In a future paper, when final morphological information is available for all of the galaxies, we will address this issue in much greater detail.

An additional element in our selection process is the rejection, from our CULL sample, of galaxies with nebular emission lines. This likely removes preferentially the later type objects from our analysis, but given that the emission selection depends in part on Hβ, it is important to test for any bias this introduces. To investigate the effects of emission selection, we have repeated the analysis *including* galaxies with emission. In this case, we find that the line-strength–σ slopes for some of the metal lines (e.g., Mg b) are on average steeper than the default solution and thus our emission selection results in a shallower metallicity gradient with σ. However, the Balmer line-strength–σ relations also steepen, suggesting that emission galaxies harbor young underlying populations not totally disguised by emission infilling. As a result, including the emission galaxies yields an age trend even stronger than in our default solution. Other surveys *corrected* their galaxies for emission (Kuntschner et al. 2001; Trager et al. 2000b) using the [O III λ5007] correction factor of Trager et al. (2000a) but still included them in their analysis of the stellar population trends. We have shown (§ 3) that such a scheme typically undercorrects at low σ and overcorrects at high σ. The net effect of this will be to *flatten* the age-σ relation derived in these studies.

The scaling relations for age, metallicity, and [α/Fe] were determined from central spectra. The fiber spectra sample a physical radius 0.2–0.8 h^{-1} kpc, depending on the distance to each cluster, and line-strength measurements were aperture corrected to a common physical radius of 0.68 h^{-1} kpc assuming universal gradients. Although the line-strength–σ relations and

the color-magnitude relation are often considered as reflecting the same underlying trends, this assumes, perhaps naively, that there is a trivial relationship between *central* stellar populations and the *global* colors.

A final set of caveats concerns the stellar population models used to translate the line-strength–σ slopes into scaling relations of population parameters. For simplicity, these models describe the highly idealized case of a single-age, single-metallicity population. In elliptical galaxies, there may instead be a broad distribution of metallicities (Harris & Harris 2002). Similarly there could be subpopulations of differing ages, as in the "frosting" models of Trager et al. (2000b). Strictly, our analysis determines scaling relations of the luminosity-weighted mean stellar age, metallicity, and [α/Fe]. An alternative to a true age trend, therefore, would be a scenario in which all galaxies have a 1 Gyr "frost" representing a progressively larger mass fraction at lower velocity dispersion. A separate concern is that the models, as described above, do not include the effects of BHB stars as observed in metal-poor globular clusters. Note, however, that the evidence for an age trend is stronger in the case of Hβ than for Hδ_F and Hγ_F. Qualitatively, this is opposite to the signature expected from variations in the BHB contribution (Schiavon et al. 2004). Robust constraints on this effect must await more sophisticated models that simultaneously incorporate BHB stars and α-element abundances.

7.2. *Comparison to Other Results on Scaling Relations*

In this section we compare our results with those from other groups, taking into account the important differences in sample selection and emission treatment as described above.

We have already noted that our results, particularly the age-σ scaling relation, are not sensitive to which Balmer line (Hδ_F, Hγ_F, Hβ) is used. This conclusion is reinforced in a separate paper (Smith 2005), which shows that a strong age gradient is also required to reproduce the slope of the NFPS Hα_A-σ relation.

Trager et al. (2000a) used principal component analysis to investigate correlations of age, metallicity, and [E/Fe] (analogous to [α/Fe]) with structural parameters in their sample of early-type galaxies in clusters and in the field. Although they model the scaling relations in a different way, fitting metallicity and [E/Fe] to age and velocity dispersion simultaneously, they find E/Fe $\propto \sigma^{0.33}$, very similar to our value of α/Fe $\propto \sigma^{0.31}$. Their metallicity gradient is dependent on the individual ages of their galaxies, but the velocity dispersion component is Z/H $\propto \sigma^{0.76}$ compared to our value of Z/H $\propto \sigma^{0.53}$. Although Trager et al. (2000a) do not claim a trend of age with σ, a simple fit to their data yields an age-σ trend with an exponent of 0.6 ± 0.2, consistent with our results.

Kuntschner et al. (2001), using line strengths for a sample drawn mostly from clusters, explicitly assume no age trend with σ, which strongly affects the other scaling relations derived. They find a higher $\Delta[Z/H]/\Delta \log \sigma$ slope ($\sim$0.9) after correcting for varying [α/Fe]. Their higher metallicity slope and lack of an age slope are likely the effect of the age-metallicity degeneracy, which we also see in the spread of points in the left panel of Figure 11. Modeling the Kuntschner et al. (2001) data according to our method, we obtain age $\propto \sigma^{\approx 0.8}$ if all three parameters are allowed to vary. Although the relation is steeper than our default solution, it is quite sensitive to a single outlying low-σ galaxy. Excluding this outlier (which has later type morphology) yields age $\propto \sigma^{\approx 0.5}$.

Poggianti et al. (2001) obtained age and metallicity estimates of $\sim$280 red galaxies in the Coma Cluster. They found a broad range in age at all magnitudes, making it difficult to quantify the mean age-σ relation, but they did note that the fraction of young dwarf galaxies in their sample is higher than the fraction of young giant galaxies. Furthermore, in Poggianti et al. (2004), the authors noted that poststarburst k+a spectra were identified in dwarf galaxy spectra, with luminosities $L \lesssim 0.1L^*$. These results indicate a "downsizing" effect in that the most recent star formation activity occurs at lower redshifts for progressively fainter galaxies. This effect is discussed further in the next section.

Caldwell et al. (2003) derived ages and metallicities for their sample of 175 early-type galaxies in clusters and in the field, including many with $\sigma < 100$ km s^{-1}. In Figure 21 of their paper, they plot age versus σ and note a strong correlation. Surprisingly, however, these authors do not quote a numerical estimate for the slope. From their Table 9 and Figure 21, we estimate their log (age)–log σ gradient to be 0.8–1.2 depending on the index combination employed. In addition, they find a shallower trend of metallicity than we do (0.32 vs. our value of 0.53). Their age gradient is steeper than ours, but their results lie on the age-metallicity degeneracy line for these two parameters.

Thomas et al. (2005) quote ages, metallicities, and α-enhancements for 54 early-type galaxies in high-density environments. They do not quote scaling relations for the sample as a whole, but rather break the data into subclasses preselected by age and velocity dispersion. In order to compare their results with ours, we have analyzed their published data for early-type galaxies in high-density environments and derive the following scaling relations from a simple unweighted regression on log σ: 0.78 ± 0.23, 0.42 ± 0.14, and 0.36 ± 0.05 for age, metallicity, and α-enhancement, respectively. These results are in excellent agreement with our results, but their errors are large because their sample contains few low-σ galaxies.

We note that Proctor et al. (2004b) also found a positive age-σ gradient in galaxies in Hickson compact groups. They do not quote an age-σ relation, but examination of their Figure 5 suggests a relation somewhat steeper ($\sim$1.2) than our best-fit value.

Figure 12 summarizes our results and those of Trager et al. (2000a), Caldwell et al. (2003), and Thomas et al. (2005). We conclude that our scaling relations of age, metallicity, and [α/Fe] fall within the range spanned by previous studies. Differences among the above results may arise in part from the different choices of line strengths used (resulting in scatter along the age-metallicity degeneracy ellipse) and from different sample selection characteristics.

7.3. *Connection to Observations at Intermediate Redshift*

The reality of our steep trend in age as a function of mass, along the red sequence, can be tested with observations at high redshift. As previously noted, the absolute age calibration of the models is less secure than the relative ages. If we identify the stellar age of the most massive systems ($\sigma \sim 400$ km s^{-1}; $\sim 20L^*$) with 13 Gyr, close to the age of the universe, then the least massive red-sequence galaxies, with $\sigma \approx 60$ km s^{-1} ($0.01L^*$), have ages of approximately 4 Gyr, and galaxies with $\sigma \approx 100$ km s^{-1} ($\sim 0.1L^*$) would have ages of $\sim$5.5 Gyr. Taking these at face value and considering also the substantial spread in age at given σ (especially for low-σ objects), we would expect to observe strong evolutionary effects at intermediate redshifts. At an earlier epoch, some of the stellar mass presently residing on the faint end of the red sequence will not have been formed at all; some will be present in star-forming galaxies, and only a fraction will exist on the red sequence itself. A key observational signature of such evolution would be a depletion or truncation of the red sequence, affecting progressively more massive galaxies with increasing look-back time.

A number of studies have in fact advanced evidence for such a depletion. For example, Smail et al. (1998) observed 10 clusters at $z \approx 0.24$, corresponding to a look-back time of 3.2 Gyr (for a concordance cosmology). Examination of their Figure 5 suggests a strong decline in the number of red-sequence galaxies at 2.0–2.5 mag below M^*. This corresponds to approximately $\sigma \sim 100$ km s^{-1}. Our age-σ relation indicates *current* ages of 3–8 Gyr for such objects, with a mean of $\sim$5 Gyr. Thus, a substantial fraction of such galaxies either had not yet become quiescent at $z \sim 0.24$ or else did not have enough time to age onto the red sequence. We conclude that the truncation of the Smail et al. (1998) red sequence is at least approximately consistent with the age-σ relation obtained from the NFPS. At a somewhat higher redshift of $z \approx 0.75$ (look-back time of 7 Gyr), De Lucia et al. (2004) examined the red sequence in clusters and found a deficiency of a factor of 2 (compared to Coma) for galaxies with $0.1L^* < L < 0.4L^*$. Our age-σ trend and scatter predict that in this luminosity range (corresponding to $\sigma \approx 100-150$ km s^{-1}), the present-day stellar age is 7 Gyr on average. Thus, we would expect $\sim$50% of such galaxies to be significantly bluer at those epochs, again approximately consistent with the observed depletion.

For a $\sim$1 Gyr period after star formation ceases, galaxies pass through a poststarburst (or E+A) phase, characterized by very strong Balmer absorption but no strong emission. Tran et al. (2003) studied E+A poststarburst galaxies in three clusters at $z = 0.3$, 0.6, and 0.8. They find that the typical velocity dispersion of poststarburst galaxies at these redshifts decreases from $\sim$170 km s^{-1} at $z = 0.8$ to $\sim$100 km s^{-1} at $z = 0.3$. If we identify these objects as galaxies in the process of fading onto the red sequence, then this trend of more massive galaxies becoming quiescent at higher redshifts fits well with our present-day age-σ trend.

In summary, our results add to recent evidence that the red sequence of cluster galaxies has built up gradually over cosmic history, progressing from more massive to less massive galaxies. Such a scenario is the cluster analog to the "downsizing" of the characteristic mass of star-forming field galaxies as discussed by Cowie et al. (1996) and Kauffmann et al. (2004).

8. CONCLUSIONS

In this paper we have presented absorption-line strength measurements for $\sim$5000 red-sequence galaxies in low-redshift clusters. Our survey samples galaxies with velocity dispersion ranging down to $\sim$50 km s^{-1}. The absorption data are complemented with emission-line measurements that can be used to select subsamples with only low levels of nebular contamination. We have employed the slopes of the line-strength–velocity dispersion relations to constrain linear scaling relations for stellar population parameters. We find that more massive galaxies are older, have higher overall metallicity, and have higher [α/Fe] ratios than galaxies of lower mass (strictly, this refers to central, luminosity-averaged properties of the stellar populations). These conclusions are quite robust and in particular are not dependent on which of the age-sensitive Balmer lines are used in the analysis. Moreover, a more traditional model-grid inversion yields consistent scaling relations.

The most important result of this paper is the very clear detection of an apparent age-mass relation for red-sequence galaxies in clusters. This arises because a flat age-σ relation cannot generate the steep slopes of the Balmer line-strength–σ relations while simultaneously matching the metal line-strength–σ slopes. This is consistent with the results from a number of other independent studies based on smaller samples.

A strong age-mass relation stands in stark contrast to the widespread assumption that cluster elliptical galaxies form a metallicity sequence of approximately constant age. Moreover, our apparently antihierarchical age-mass relation disagrees with the predictions from semianalytic galaxy formation models, which suggest either that brighter elliptical galaxies have slightly younger stellar populations (Kauffmann & Charlot 1998) or that early-type cluster galaxies should have uniformly old stellar populations (Fig. 18 of Kuntschner et al. 2002, which is based on the models of Cole et al. 2000).

Our results are broadly consistent with claims of a truncated or depleted red sequence in clusters at higher redshifts (e.g., Smail et al. 1998; De Lucia et al. 2004) and of an increase in the average masses of poststarburst cluster members with increasing redshift (Tran et al. 2003). Together, these observations suggest a trend of downsizing galaxy formation in clusters, mirroring a similar decline in the characteristic mass of star-forming field galaxies (Cowie et al. 1996; Kauffmann et al. 2004). These studies all present a picture of the red sequence in clusters building up slowly over cosmic history, proceeding from the most massive to progressively lower mass galaxies. In particular, galaxies presently on the faint end of the red sequence became quiescent only at very recent epochs and are likely the descendants of star-forming or poststarburst galaxies in intermediate-redshift clusters.

We gratefully acknowledge the substantial assignment of NOAO observing resources to the NFPS program and a Chretien International Research Grant from the American Astronomical Society. J. E. N. was supported by a Dartmouth College Fellowship, a Graduate Assistance in Areas of National Need fellowship, and a NASA space grant. M. J. H. acknowledges support from the NSERC of Canada, an Ontario Premier's Research Excellence Award. S. A. W. M. was supported from the PPARC grant "Extragalactic Astronomy & Cosmology at Durham 1998–2002." S. J. Q. was supported by a PPARC studentship. IRAF is distributed by the National Optical Astronomy Observatory, which is operated by the Association of Universities for Research in Astronomy, Inc., under contract with the National Science Foundation.

REFERENCES

Abazajian, K., et al. 2004, AJ, 128, 502
Balogh, M. L., Morris, S. L., Yee, H. K. C., Carlberg, R. G., & Ellingson, E. 1999, ApJ, 527, 54
Bender, R., Burstein, D., & Faber, S. M. 1993, ApJ, 411, 153
Bernardi, M., et al. 2003, AJ, 125, 1882
Bower, R. G., Lucey, J. R., & Ellis, R. S. 1992, MNRAS, 254, 601
Burstein, D., Faber, S. M., Gaskell, C. M., & Krumm, N. 1984, ApJ, 287, 586
Caldwell, N., Rose, J. A., & Concannon, K. D. 2003, AJ, 125, 2891
Cardiel, N., Gorgas, J., Cenarro, J., & Gonzalez, J. J. 1998, A&AS, 127, 597
Cenarro, A. J., Cardiel, N., Gorgas, J., Peletier, R. F., Vazdekis, A., & Prada, F. 2001, MNRAS, 326, 959
Cole, S., Lacey, C. G., Baugh, C. M., & Frenk, C. S. 2000, MNRAS, 319, 168
Colless, M., Burstein, D., Davies, R. L., McMahan, R. K., Saglia, R. P., & Wegner, G. 1999, MNRAS, 303, 813
Cowie, L. L., Songaila, A., Hu, E. M., & Cohen, J. G. 1996, AJ, 112, 839
Davies, R. L., Burstein, D., Dressler, A., Faber, S. M., Lynden-Bell, D., Terlevich, R. J., & Wegner, G. 1987, ApJS, 64, 581
De Lucia, G., et al. 2004, ApJ, 610, L77
Djorgovski, S., & Davis, M. 1987, ApJ, 313, 59
Dressler, A., Lynden-Bell, D., Burstein, D., Davies, R. L., Faber, S. M., Terlevich, R. J., & Wegner, G. 1987, ApJ, 317, 42

Ebeling, H., Edge, A. C., Bohringer, H., Allen, S. W., Crawford, C. S., Fabian, A. C., Voges, W., & Huchra, J. P. 1998, MNRAS, 301, 881
Ebeling, H., Voges, W., Bohringer, H., Edge, A. C., Huchra, J. P., & Briel, U. G. 1996, MNRAS, 281, 799
Faber, S. M., & Jackson, R. E. 1976, ApJ, 204, 668
González, J. J. 1993, Ph.D. thesis, Univ. California, Santa Cruz
Goudfrooij, P., de Jong, T., Hansen, L., & Norgaard-Nielsen, H. U. 1994, MNRAS, 271, 833
Harris, W. E., & Harris, G. L. H. 2002, AJ, 123, 3108
Herbig, G. H., & Mayall, N. U. 1957, PASP, 69, 563
Jorgensen, I. 1997, MNRAS, 288, 161
Jorgensen, I., Franx, M., & Kjaergaard, P. 1995, MNRAS, 276, 1341
Kauffmann, G., & Charlot, S. 1998, MNRAS, 294, 705
Kauffmann, G., White, S. D. M., Heckman, T. M., Ménard, B., Brinchmann, J., Charlot, S., Tremonti, C., & Brinkmann, J. 2004, MNRAS, 353, 713
Kinney, A. L., Calzetti, D., Bohlin, R. C., McQuade, K., Storchi-Bergmann, T., & Schmitt, H. R. 1996, ApJ, 467, 38
Kuntschner, H., Lucey, J. R., Smith, R. J., Hudson, M. J., & Davies, R. L. 2001, MNRAS, 323, 615
Kuntschner, H., Smith, R. J., Colless, M., Davies, R. L., Kaldare, R., & Vazdekis, A. 2002, MNRAS, 337, 172
Maraston, C. 1998, MNRAS, 300, 872
Maraston, C., & Thomas, D. 2000, ApJ, 541, 126
Mehlert, D., Thomas, D., Saglia, R. P., Bender, R., & Wegner, G. 2003, A&A, 407, 423
Moore, S. A. W., Lucey, J. R., Kuntschner, H., & Colless, M. 2002, MNRAS, 336, 382
Phillips, M. M., Jenkins, C. R., Dopita, M. A., Sadler, E. M., & Binette, L. 1986, AJ, 91, 1062
Poggianti, B. M., Bridges, T. J., Komiyama, Y., Yagi, M., Carter, D., Mobasher, B., Okamura, S., & Kashikawa, N. 2004, ApJ, 601, 197
Poggianti, B. M., et al. 2001, ApJ, 562, 689
Proctor, R. N. 2002, Ph.D. thesis, Univ. Central Lancashire
Proctor, R. N., Forbes, D. A., & Beasley, M. A. 2004a, MNRAS, 355, 1327
Proctor, R. N., Forbes, D. A., Hau, G. K. T., Beasley, M. A., De Silva, G. M., Contreras, R., & Terlevich, A. I. 2004b, MNRAS, 349, 1381
Saglia, R. P., Maraston, C., Thomas, D., Bender, R., & Colless, M. 2002, ApJ, 579, L13
Sandage, A., & Visvanathan, N. 1978, ApJ, 225, 742
Schiavon, R. P., Rose, J. A., Courteau, S., & MacArthur, L. A. 2004, ApJ, 608, L33
Simard, L., et al. 2002, ApJS, 142, 1
Smail, I., Edge, A. C., Ellis, R. S., & Blandford, R. D. 1998, MNRAS, 293, 124
Smith, R. J. 2005, MNRAS, 359, 975
Smith, R. J., et al. 2004, AJ, 128, 1558 (Paper I)
Stasińska, G., Mateus, A., Sodré, L., & Szczerba, R. 2004, A&A, 420, 475
Terlevich, A. I., & Forbes, D. A. 2002, MNRAS, 330, 547
Thomas, D., Maraston, C., & Bender, R. 2003, MNRAS, 339, 897 (TMB03)
Thomas, D., Maraston, C., Bender, R., & Mendes de Oliveira, C. 2005, ApJ, 621, 673
Thomas, D., Maraston, C., & Korn, A. 2004, MNRAS, 351, L19
Trager, S. C., Faber, S. M., Worthey, G., & González, J. J. 2000a, AJ, 119, 1645
———. 2000b, AJ, 120, 165
Trager, S. C., Worthey, G., Faber, S. M., Burstein, D., & Gonzalez, J. J. 1998, ApJS, 116, 1
Tran, K. H., Franx, M., Illingworth, G., Kelson, D. D., & van Dokkum, P. 2003, ApJ, 599, 865
van Dokkum, P. G., & Ellis, R. S. 2003, ApJ, 592, L53
Vazdekis, A. 1999, ApJ, 513, 224
Wegner, G., Colless, M., Saglia, R. P., McMahan, R. K., Davies, R. L., Burstein, D., & Baggley, G. 1999, MNRAS, 305, 259
Worthey, G. 1994, ApJS, 95, 107
Worthey, G., & Ottaviani, D. L. 1997, ApJS, 111, 377
Wuyts, S., van Dokkum, P. G., Kelson, D. D., Franx, M., & Illingworth, G. D. 2004, ApJ, 605, 677